

Aus der Klinik für Innere Medizin II
Universitätsklinikum des Saarlandes, Homburg/Saar

**Entwicklung eines Prognosemodells für die Dekompensation der Leberzirrhose mithilfe
Künstlicher Intelligenz/Maschinellen Lernen**

Dissertation zur Erlangung des Grades eines Doktors der Medizin

der Medizinischen Fakultät

der UNIVERSITÄT DES SAARLANDES

2024

vorgelegt von: Sophie Elisabeth Müller
geb. am 28.03.2000 in Homburg/Saar

Tag der Promotion: 02.10.2025

Dekan: Univ.-Prof. Dr. med. dent. Matthias Hannig

Berichterstatter: Univ.-Prof. Dr. med. Frank Lammert

Prof. Dr. Olga Kalinina

Inhaltsverzeichnis

Abbildungsverzeichnis.....	VII
Tabellenverzeichnis.....	IX
Abkürzungsverzeichnis.....	XI
1 Zusammenfassung.....	1
1.1 Deutsche Zusammenfassung.....	1
1.2 Abstract.....	2
2 Einleitung.....	3
2.1 Zirrhose.....	3
2.1.1 Stadien und natürlicher Verlauf der Leberzirrhose.....	4
2.1.2 Komplikationen der Leberzirrhose.....	5
2.1.2.1 Portale Hypertension.....	5
2.1.2.2 Komplikationen durch die Hepatozytendysfunktion.....	6
2.1.2.3 Infektionen bei Leberzirrhose.....	7
2.1.2.3.1 Spontan bakterielle Peritonitis.....	7
2.1.2.3.2 <i>NOD2</i> -Gen.....	8
2.1.3 Prognose der Zirrhose und deren Abschätzung.....	8
2.2 Künstliche Intelligenz, Machine Learning, Deep Learning.....	11
2.3 Künstliche Intelligenz und maschinelles Lernen in der Hepatologie.....	12
2.4 Ziel der Studie.....	13
3 Patienten und Methodik.....	14
3.1 Patienten.....	14
3.2 Methodik.....	19
3.2.1 Grundlagen.....	19
3.2.1.1 Machine Learning-Verfahren.....	19
3.2.1.1.1 Entscheidungsbäume.....	19
3.2.1.1.2 Random Forests.....	20
3.2.1.1.3 Support Vector Machines.....	21

3.2.1.1.4	Neuronale Netze.....	22
3.2.1.2	Inspektion des Modells: Permutation Feature Importance.....	25
3.2.1.3	Metriken zur Evaluation der Verfahren	26
3.2.2	Umsetzung/Implementierung	27
3.2.2.1	Vorbereitung der Daten.....	27
3.2.2.2	Clusteranalyse	28
3.2.2.3	Ablauf des Trainierens und Testens.....	30
4	Ergebnisse	32
4.1	Entscheidungsbäume.....	33
4.1.1	Retrospektive Beurteilung.....	33
4.1.2	Prospektive Beurteilung.....	35
4.2	Random Forests.....	37
4.2.1	Retrospektive Beurteilung.....	37
4.2.2	Prospektive Beurteilung.....	39
4.3	Support Vector Machine	41
4.3.1	Retrospektive Beurteilung.....	41
4.3.2	Prospektive Beurteilung.....	43
4.4	Neuronale Netze.....	45
4.4.1	Retrospektive Beurteilung.....	45
4.4.2	Prospektive Beurteilung.....	47
4.5	Vergleich der Ergebnisse	49
5	Diskussion	53
5.1	Zusammenfassung der Ergebnisse	53
5.2	Vergleich der Ergebnisse der einzelnen Verfahren.....	53
5.3	Diskussion der Parameter.....	54
5.4	Aktuelle Studienlage bezüglich Dekompensation	55
5.5	Vergleich der Ergebnisse mit bestehenden Scores.....	56
5.6	Einsatz von maschinellem Lernen	58
5.7	Stärken der Studie	59

5.8	Limitationen der Studie.....	59
5.9	Zusammenfassung und Ausblick	60
6	Literaturverzeichnis.....	61
7	Publikationen.....	70
	Dank	73
	Lebenslauf.....	75

Abbildungsverzeichnis

Abbildung 1: Stadien der Zirrhose.....	4
Abbildung 2: Dendrogramm Clusteranalyse.....	29
Abbildung 3: Top 20 Entscheidungsbaum ungeclustert retrospektiv	33
Abbildung 4: Top 20 Entscheidungsbaum geclustert retrospektiv	34
Abbildung 5: Top 20 Entscheidungsbaum ungeclustert prospektiv	35
Abbildung 6: Top 20 Entscheidungsbaum geclustert prospektiv	36
Abbildung 7: Top 20 Random Forest ungeclustert retrospektiv.....	37
Abbildung 8: Top 20 Random Forest geclustert retrospektiv	38
Abbildung 9: Top 20 Random Forest ungeclustert prospektiv	39
Abbildung 10: Top 20 Random Forest geclustert prospektiv	40
Abbildung 11: Top 20 SVM ungeclustert retrospektiv.....	41
Abbildung 12: Top 20 SVM geclustert retrospektiv.....	42
Abbildung 13: Top 20 SVM ungeclustert prospektiv	43
Abbildung 14: Top 20 SVM geclustert prospektiv	44
Abbildung 15: Top 20 neuronales Netz ungeclustert retrospektiv.....	45
Abbildung 16: Top 20 neuronales Netz retrospektiv geclustert.....	46
Abbildung 17: Top 20 neuronales Netz ungeclustert prospektiv.....	47
Abbildung 18: Top 20 neuronales Netz geclustert prospektiv.....	48

Tabellenverzeichnis

Tabelle 1: Child-Pugh Score	9
Tabelle 2: Charakteristika der kategorialen Variablen.....	16
Tabelle 3: Charakteristika der numerischen Variablen.....	18
Tabelle 4: Accuracy der einzelnen Verfahren zur retrospektiven Beurteilung einer Dekompensation.....	32
Tabelle 5: Accuracy bezüglich der prospektiven Einschätzung einer Dekompensation	32
Tabelle 6: Metriken Entscheidungsbaum ungeclustert retrospektiv	33
Tabelle 7: Metriken Entscheidungsbaum geclustert retrospektiv	34
Tabelle 8: Metriken Entscheidungsbaum ungeclustert prospektiv	35
Tabelle 9: Metriken Entscheidungsbaum geclustert prospektiv	36
Tabelle 10: Metriken Random Forest ungeclustert retrospektiv.....	37
Tabelle 11: Metriken Random Forest geclustert retrospektiv.....	38
Tabelle 12: Metriken Random Forest ungeclustert prospektiv	39
Tabelle 13: Metriken Random Forest geclustert prospektiv	40
Tabelle 14: Metriken SVM ungeclustert retrospektiv.....	41
Tabelle 15: Metriken SVM geclustert retrospektiv.....	42
Tabelle 16: Metriken SVM ungeclustert prospektiv	43
Tabelle 17: Metriken SVM geclustert prospektiv	44
Tabelle 18: Metriken neuronales Netz ungeclustert retrospektiv	45
Tabelle 19: Metriken neuronales Netz geclustert retrospektiv	46
Tabelle 20: Metriken neuronales Netz ungeclustert prospektiv.....	47
Tabelle 21: Metriken neuronales Netz geclustert prospektiv.....	48
Tabelle 22: Vergleich Top 20 retrospektiv ungeclustert.....	49
Tabelle 23: Vergleich Top 20 retrospektiv geclustert.....	50
Tabelle 24: Vergleich Top 20 prospektiv ungeclustert.....	51
Tabelle 25: Vergleich Top 20 prospektiv geclustert.....	52
Tabelle 26: Entscheidungsbaum ungeclustert retrospektiv.....	77
Tabelle 27: Entscheidungsbaum geclustert retrospektiv.....	79
Tabelle 28: Entscheidungsbaum ungeclustert prospektiv	81
Tabelle 29: Entscheidungsbaum geclustert prospektiv	83
Tabelle 30: Random Forest ungeclustert retrospektiv	84
Tabelle 31: Random Forest geclustert retrospektiv	86
Tabelle 32: Random Forest ungeclustert prospektiv.....	88
Tabelle 33: Random Forest geclustert prospektiv.....	90
Tabelle 34: SVM ungeclustert retrospektiv	91
Tabelle 35: SVM geclustert retrospektiv	93
Tabelle 36: SVM ungeclustert prospektiv.....	94

Tabelle 37: SVM geclustert prospektiv.....	96
Tabelle 38: Neuronales Netz ungeclustert retrospektiv	97
Tabelle 39: Neuronales Netz geclustert retrospektiv	99
Tabelle 40: Neuronales Netz ungeclustert prospektiv	101
Tabelle 41: Neuronales Netz geclustert prospektiv	103

Abkürzungsverzeichnis

ACLF	akut-auf-chronisches Leberversagen
AFP	Alpha-Fetoprotein
AI	Artificial Intelligence
ALBI score	Albumin-Bilirubin score
ALT	Alanin-Aminotransferase
AP	Alkalische Phosphatase
AST	Aspartat-Aminotransferase
AUC	area under the curve
bzw.	beziehungsweise
Ca	Calcium
CHE	Cholinesterase
CLIF-C AD	Chronic Liver Failure Consortium Acute Decompensation
CNN	Convolutional Neural Network
CRP	C-reaktives Protein
CT	Computertomographie
d. h.	das heißt
DT	Decision Tree, Entscheidungsbaum
EPOD	Early prediction of decompensation
FIB-4	Fibrosis-4
GGT	γ -Glutamyltransferase
HB	Hämoglobin
HBV	Hepatitis B-Virus
HCC	hepatozelluläres Karzinom
HCV	Hepatitis C-Virus
HE	hepatische Enzephalopathie
HVPG	hepatisch-venöser Druckgradient
HRS	hepatorenales Syndrom
i. d. R.	in der Regel
INCA	Impact of <i>NOD2</i> genotype-guided antibiotic prevention on survival in patients with liver Cirrhosis and Ascites
IG	Informationsgewinn

INR	International Normalized Ratio
K	Kalium
KI	Künstliche Intelligenz
KNN	künstliches Neuronales Netz
LSM	Liver Stiffness Measurement (Lebersteifigkeitsmessung)
Max	Maximum
MELD	Model for End-Stage Liver Disease
MELD-Na	Model for End-Stage Liver Disease-Natrium
Min	Minimum
ML	maschinelles Lernen, Machine Learning
MRT	Magnetresonanztomographie
Na	Natrium
NAFLD	nichtalkoholische Fettlebererkrankung
NASH	nichtalkoholische Steatosis hepatis
<i>NOD2</i>	nucleotide-binding oligomerisation domain 2
PBC	primär biliäre Cholangitis
PFI	Permutation Feature Importance
PPI	Protonenpumpeninhibitor
PSC	primär sklerosierende Cholangitis
RFC	Random Forest Classifier
ROC	receiver operating characteristic
SPB	spontan bakterielle Peritonitis
SVM	Support Vector Maschine
TAG	Triacylglyceride
u. a.	unter anderem
v. a.	vor allem
z. B.	zum Beispiel

1 Zusammenfassung

1.1 Deutsche Zusammenfassung

Einleitung und Ziel: Bei Patienten mit Zirrhose erhöhen Dekompensationen die Mortalität signifikant, weshalb eine frühe Erkennung von Risikopatienten von entscheidender Bedeutung ist. Diese Arbeit hatte zum Ziel, mithilfe maschineller Lernverfahren Parameter zu bestimmen, die Dekompensationen vorhersagen können.

Patienten und Methoden: Basis der Studie bildete die Datenbank der INCA-Studie (Impact of nucleotide-binding oligomerization domain-containing protein 2 [*NOD2*] genotype-guided antibiotic prevention on survival in patients with liver Cirrhosis and Ascites), welche retrospektive und prospektive Daten von 983 Patienten mit Zirrhose aus den Universitätskliniken Homburg und Halle enthält. Zusätzlich zu klinischen Informationen wurden auch Laborwerte und *NOD2*-Genotypen betrachtet. Für das maschinelle Lernen wurden Entscheidungsbäume, Random Forests, Support Vector Maschinen und neuronale Netze genutzt. Zur Identifizierung der für die Vorhersage wichtigen Parameter wurde zunächst eine hierarchische Clusteranalyse basierend auf der Spearman-Rangkorrelation und der Ward-Minimumvarianzmethode durchgeführt, anschließend wurde als Modellinspektionstechnik die Permutation Feature Importance eingesetzt.

Ergebnisse: Von den 983 Patienten wiesen 313 Patienten zum Indexzeitpunkt eine kompensierte Zirrhose auf, 354 Patienten hatten in der Vergangenheit eine Dekompensation erlitten, waren jedoch zum Inklusionszeitpunkt re kompensiert, und 316 Patienten befanden sich im de kompensierten Zustand. Am Follow-up nahmen 825 Patienten teil, dieses hatte eine mittlere Dauer von zwölf Monaten und eine maximale Dauer von 55 Monaten. 46,5 % der Patienten de kompensierten. Mit einer Accuracy von 81,6 % auf den Trainings- und 70,5 % auf den Testdaten erzielten Random Forests bei der retrospektiven Vorhersage einer Dekompensation die besten Ergebnisse. Support Vector Maschinen konnten Dekompensationen am besten prospektiv vorhersagen, sie erreichten eine Accuracy von 78,6 % bzw. 73,8 %. Die Analyse der Permutation Feature Importance ergab, dass die Ausgangswerte von Bilirubin und Albumin sowie der maximale Bilirubinwert für die retrospektive Analyse entscheidend waren. Bei der prospektiven Analyse wurden der maximale Bilirubinwert und die Ausgangswerte von Natrium und Albumin am höchsten gewichtet. Neben Parametern etablierter Risikoscores wie dem Child-Pugh- und dem Model for End-stage Liver Disease-Score, waren Entzündungswerte und *NOD2*-Genotypen von hoher Relevanz.

Schlussfolgerungen: Unter den getesteten Modellen erzielten Random Forests und Support Vector Maschinen die höchste Genauigkeit für die Vorhersage einer Dekompensation. Neben Laborwerten wie Bilirubin und Albumin, spielten Infektionen und genetische Faktoren eine entscheidende Rolle für die individuelle Vorhersage.

1.2 Abstract: Machine learning models predicting decompensation in cirrhosis

Background and Aims: Decompensation of cirrhosis significantly increases the mortality rate, thus an early detection of patients at risk is of utmost importance. This study aimed to identify parameters which predict a decompensation using machine learning.

Methods: Several machine learning techniques including Decision Trees, Random Forests, Support Vector Machines, and Neural Networks were applied to the INCA trial database (Impact of nucleotide-binding oligomerization domain-containing protein 2 [*NOD2*] genotype-guided antibiotic prevention on survival in patients with liver Cirrhosis and Ascites) containing both prospective and retrospective data from 983 patients with cirrhosis from the university hospitals Homburg and Halle. Besides laboratory values, clinical information and *NOD2* genotypes were considered. After performing a hierarchical clustering on the numerical variables based on the Spearman rank-order correlation using Ward's minimum variance method, Permutation Feature Importance was used as model inspecting technique to determine the impact of the parameters on the prediction of a decompensation.

Results: Of the 983 patients, 313 patients had compensated cirrhosis at the index date, 354 patients were decompensated earlier, but currently recompensated, and 316 were in a decompensated state. Overall, 825 patients were included in the follow-up (median duration: 12 months, maximum: 55 months), of whom 46.5 % decompensated. In the retrospective assessment, Random Forests performed best with an accuracy of 81.6 % on training and 70.5 % on test data. Support Vector Machines were best at predicting a decompensation prospectively, achieving an accuracy of 78.6 % and 73.8 %, respectively. Permutation Feature Importance demonstrated that the baseline values of albumin and bilirubin and the maximum bilirubin value were the highest ranked parameters to be associated with a former decompensation. For the prospective analysis, the maximum bilirubin value and the baseline values of albumin and sodium were the highest ranked. Alongside parameters of established risk scores like the Child-Pugh and the Model for End-stage Liver Disease score, parameters related to infections (e.g., inflammatory markers and intake of antibiotics) and *NOD2* genotypes were highly ranked.

Conclusion:

Among tested machine learning models, Random Forests and Support Vector Machines seem to have the highest accuracy in predicting decompensation. In addition to laboratory parameters, *NOD2* genotypes and infections were crucial parameters for the individual prediction.

2 Einleitung

2.1 Zirrhose

Die Zirrhose wird als irreversible Folge der meisten chronischen Lebererkrankungen betrachtet. Eine chronische Leberzellschädigung führt zu einer zunehmenden fibrösen Transformation und Destruktion des Parenchyms, wodurch einerseits die Synthese- und Entgiftungsfunktion eingeschränkt wird und andererseits eine portale Hypertension mit deren Komplikationen auftreten kann.

Mit einer Inzidenz von etwa 250/100.000 Einwohner pro Jahr und ca. 300.000-400.000 Erkrankten in Deutschland ist die Leberzirrhose in Industrieländern eine recht häufige Erkrankung [49]. Am häufigsten entsteht sie infolge eines Alkoholmissbrauchs (50 %), gefolgt von chronischen Virushepatitiden (insbesondere durch das Hepatitis-B- und Hepatitis-C-Virus (HBV bzw. HCV), 30 %) und einer nicht-alkoholischen Fettlebererkrankung (NASH, ca. 10 %) [88]. Seltener Ursachen stellen z. B. die primär biliäre oder primär sklerosierende Cholangitis (PBC, PSC), die Autoimmunhepatitis und Stoffwechselerkrankungen wie α 1-Antitrypsinmangel, Hämochromatose und Morbus Wilson dar. Auch infolge einer chronischen Stauung oder Gefäßanomalien kann eine Zirrhose entstehen. Bei bis zu 10 % der Erkrankten lässt sich keine Ursache feststellen (kryptogene Zirrhose).

Durch den Einfluss schädigender Noxen werden Hepatozytennekrosen und inflammatorische Reaktionen ausgelöst, wodurch ortsständige Makrophagen, die Kupffer-Zellen, aktiviert werden. Diese wiederum aktivieren über weitere proinflammatorische Zytokine die Itozellen, welche mit einer übermäßigen Synthese von extrazellulären Matrixbestandteilen reagieren (Leberfibrogenese) [114]. Zudem kommt es zu einer Einwanderung und Aktivierung von Monozyten und Granulozyten aus der Blutbahn. Infolge überschießender bindegewebiger Reparaturvorgänge werden u. a. vaskularisierte fibröse portoportale und portozentrale Septen ausgebildet, die über eine intrahepatische portosystemische Shuntbildung zur Minderversorgung der Hepatozyten bzw. Minderdurchblutung der Leber beitragen. Durch eine massive Einlagerung von extrazellulären Matrixbestandteilen in den Dissé-Raum geht die Fenestrierung des sinusoidalen Endothels verloren, was zu einer gestörten Clearance insbesondere makromolekularer Stoffe und zu einem erhöhten Flusswiderstand führt. Im Verlauf entsteht eine gestörte Histoarchitektur mit einer ungeordneten Ausbildung von Regeneratknoten.

Im Gegensatz zur Zirrhose mit Veränderungen der Gefäßarchitektur ist die reine Fibrose der Leber noch reversibel.

2.1.1 Stadien und natürlicher Verlauf der Leberzirrhose

Der klinische Verlauf einer Zirrhose wird nach D'Amico in vier Stadien unterteilt [32]. Dabei unterscheidet man grundsätzlich zwischen dem initialen, kompensierten Stadium, in dem die Patienten meist asymptomatisch sind und ein normaler oder nur leicht erhöhter portaler Druck vorliegt (Stadium 1 und 2), und dem symptomatischen, dekompenzierten Stadium (Stadium 3 und 4). Eine Dekompensation liegt vor, wenn Aszites, Ikterus, Varizenblutungen oder eine hepatische Enzephalopathie auftreten [44,78], wobei Aszites meist als erstes klinisch apparent wird.

Das Stadium 1 zeichnet sich dadurch aus, dass die Zirrhose kompensiert ist und noch keine Ösophagusvarizen vorliegen. Im Stadium 2 ist die Zirrhose ebenfalls kompensiert, allerdings liegen hier Ösophagusvarizen (ohne stattgehabte Blutung) vor. Mit dem Auftreten von Aszites erfolgt der Übergang in das Stadium 3, eine Ösophagusvarizenblutung definiert das Stadium 4 (unabhängig vom Vorhandensein von Aszites).

Schätzungsweise erfolgt bei 5-7 % der Patienten pro Jahr der Übergang von einer kompensierten in eine dekompenzierte Zirrhose [31], womit auch ein starker Anstieg des Mortalitätsrisikos einhergeht. Im dekompenzierten Stadium beträgt das mediane Überleben lediglich zwei Jahre im Vergleich zu über zwölf Jahren im kompensierten Stadium [32].

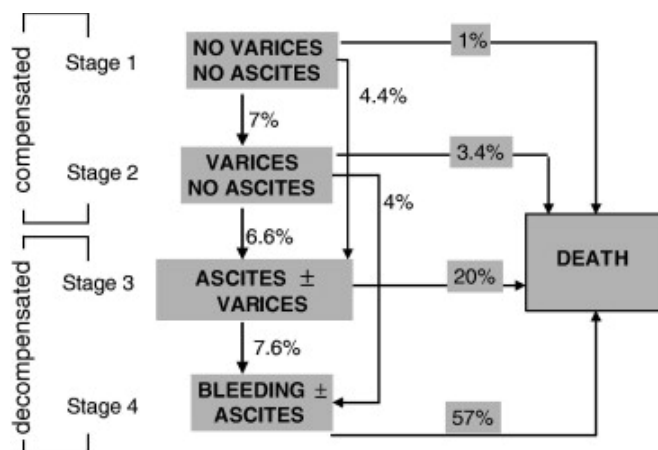


Abbildung 1: Stadien der Zirrhose, entnommen aus [32]

2.1.2 Komplikationen der Leberzirrhose

Grundsätzlich kann man bei einer Zirrhose zwischen Komplikationen bedingt durch eine portale Hypertension und Komplikationen bedingt durch die verminderte metabolische bzw. exokrine Funktion der Leber differenzieren.

2.1.2.1 Portale Hypertension

Normalerweise beträgt der portalvenöse Druck 3-6 mmHg, ab einem Pfortaderdruck > 10 mmHg liegt eine klinisch relevante portale Hypertonie vor [12,48].

Dabei kann man zwischen prä-, intra- und posthepatischen Ursachen einer portalen Hypertension unterscheiden. Zu den prähepatischen Ursachen zählen z. B. Thrombosen von Pfortader oder Milzvene, zu den Posthepatischen u. a. das Budd-Chiari-Syndrom oder eine Rechtsherzinsuffizienz. Die Leberzirrhose als intrahepatische Ursache ist in Europa führend, bis zu 90 % der asymptomatischen Leberzirrotiker zeigen einen erhöhten portalvenösen Druck [12].

Die im Rahmen einer Leberzirrhose auftretende portale Hypertonie beruht nicht nur auf der statischen Veränderung der Mikrozirkulation durch Fibrosierung, Angiogenese und Gefäßobstruktion, sondern auch auf einem erhöhten intrahepatischen Gefäßtonus durch die Kontraktion aktivierter Itozellen und Myofibroblasten in den Sinusioden und fibrösen Septen sowie von Myozyten im hepatischen Gefäßsystem [67,68,99,110]. Dies wird durch eine erhöhte lokale Produktion von Vasokonstriktoren (u. a. Angiotensin II, Endothelin-1, Thromboxan A2 und Noradrenalin) sowie verminderte Bioverfügbarkeit von Vasodilatoren wie Stickstoffmonoxid bedingt und durch eine überschießende Reaktion der Itozellen zusätzlich begünstigt [12,43].

Im weiteren Krankheitsverlauf führt eine gesteigerte Sekretion von Vasodilatoren (u. a. Stickstoffmonoxid [118], Prostazykline, Endocannabinoide) und gestörte Reaktion auf Vasokonstriktoren zu einer ausgeprägten arteriölen Vasodilatation und Neoangiogenese im Splanchnikusgebiet [15,43,58]. Durch ein vaskuläres „Underfilling“ mit einer systemischen Hypotonie [106] werden vasoaktive endogene Systeme stimuliert, was über eine Aktivierung des Sympathikus und des Renin-Angiotensin-Aldosteron-Systems sowie über eine Ausschüttung des antidiuretischen Hormons eine Natrium- und Wasserretention verursacht. Das so retinierte Wasser sammelt sich u. a. als Aszites in der Peritonealhöhle. Verstärkt wird die Aszitesbildung durch eine gesteigerte Lymphproduktion durch die erhöhte Gefäßpermeabilität [22] und durch einen Abfall des kolloidosmotischen Drucks im Rahmen einer Hypoalbuminämie bei reduzierter Lebersyntheseleistung [34].

Zudem begünstigt die arterioläre Vasodilatation zusammen mit einem erhöhten kardialen Auswurf die Ausbildung portosystemischer Kollateralkreisläufe [58]. Dies wird u. a. über proangiogenetische Faktoren wie vascular endothelial growth factor vermittelt [1,83], sodass präexistente Gefäße wiedereröffnet und neue Gefäße gebildet werden.

Durch die Ausbildung portokavaler Umgehungskreisläufe können Varizen an multiplen Stellen entstehen. Am bedrohlichsten sind Ösophagus- und Magenfundusvarizen aufgrund der hohen Blutungsgefahr, es können aber auch anorektale Varizen sowie Bauchhautvarizen auftreten. Durch die Dilatation von Magenschleimhautgefäßen entsteht eine portal-hypertensive Gastropathie.

Im Rahmen des vaskulären „Underfilling“ bei Aszites kommt es in der Niere kompensatorisch zu einer Vasokonstriktion und verstärkten Retention von Wasser und Natrium, um den Perfusionsdruck zu erhalten. Insbesondere getriggert durch hohe Volumenverluste kann die glomeruläre Filtrationsrate abnehmen und die Niere geschädigt werden (hepatorenales Syndrom).

In der Lunge fällt aufgrund einer gesteigerten Perfusion und intrapulmonalen Vasodilatation der arterielle Sauerstoffgehalt ab (portopulmonales Syndrom). Das kardiovaskuläre System reagiert auf den intravasalen Flüssigkeitsmangel mit einer Tachykardie, zudem wird durch die portale Hypertension eine systolische und diastolische Funktionsstörung mit einem Remodeling begünstigt.

Zudem kann über den Rückstau von Blut in die Milz im Rahmen einer portalen Hypertension eine Splenomegalie mit Hypersplenismus und Panzytopenie, insbesondere Thrombozytopenie, ausgelöst werden.

2.1.2.2 Komplikationen durch die Hepatozytendysfunktion

Im Rahmen einer Hepatozytendysfunktion ist der Bilirubinabbau gestört. Dies kann sich klinisch durch einen Ikterus äußern, dieser ist definiert durch eine Serumbilirubinkonzentration ≥ 3 mg/dl oder einen Bilirubinanstieg um mindestens 3 mg/dl.

Zudem kann der Abbau neurotoxischer Stoffe wie Ammoniak, Mercaptane und γ -Aminobuttersäure gestört sein, sodass diese akkumulieren. Ammoniak passiert die Blut-Hirn-Schranke und wird von den Astrozyten metabolisiert. Aufgrund osmotischer Effekte schwellen die Astrozyten an, eine hepatische Enzephalopathie (HE) als potentiell reversible Funktionsstörung des Gehirns entsteht [39]. Die HE ist gekennzeichnet durch diverse neurologische und psychiatrische Auffälligkeiten und wird gemäß den West-Haven-Kriterien in vier Stadien eingeteilt. Diese reichen von einer milden Funktionsstörung mit Konzentrationsproblemen und Stimmungsschwankungen im Stadium 1 bis zum Coma hepaticum im Stadium 4 [39,78,112].

2.1.2.3 Infektionen bei Leberzirrhose

Bakterielle Infektionen treten sehr häufig im Rahmen einer Leberzirrhose auf und sind eine der Hauptursachen für das Fortschreiten des Leberversagens, für die Entwicklung von Komplikationen und für eine erhöhte Mortalität [59]. So konnten Studien zeigen, dass die Mortalität durch bakterielle Infektionen etwa um das Vierfache erhöht wird [5,42]. Daher wird auch von manchen Experten die Ansicht vertreten, dass bakterielle Infektionen ein eigenständiges prognostisches Stadium (Stadium 5) darstellen [38].

Infektionen und Dekompensationen beeinflussen sich wechselseitig: Einerseits kommt es im Rahmen bakterieller Infektionen häufiger zu einer Dekompensation, andererseits treten bei einer Dekompensation auch vermehrt Infektionen auf [24,38,40,79]. Sowohl die Inzidenz als auch die Schwere von Infektionen steigt mit der Abnahme der Leberfunktion [14,40].

Harnwegsinfektionen und die spontan-bakterielle Peritonitis (SBP) stellen die häufigsten bakteriellen Infektionen bei Leberzirrhose dar, gefolgt von Pneumonien, Weichteilinfektionen und spontanen Bakteriämien [59].

2.1.2.3.1 Spontan bakterielle Peritonitis

Die SBP ist mit etwa 24 % die häufigste bakterielle Infektion bei hospitalisierten Patienten mit Zirrhose und Aszites [40].

Bei der SBP handelt es sich um eine primäre Peritonitis ohne Hinweis auf einen anderen intestinalen Fokus (wie z. B. Cholezystitis oder Divertikulitis), welche im ambulanten Bereich meist monobakteriell durch gramnegative Keime des Gastrointestinaltraktes (insbesondere *Escherichia coli*) verursacht wird, seltener spielen grampositive Bakterien wie Streptokokken oder Pilze eine Rolle. Im nosokomialen Umfeld dominieren die grampositiven Erreger [40].

Definiert ist eine SBP durch den Nachweis von über 250 neutrophilen Granulozyten pro mm³ Aszites [45]. Gelingt der mikrobiologische Nachweis einer bakteriellen Kolonisation des Aszites, ohne dass dabei eine erhöhte Anzahl neutrophiler Granulozyten vorliegt, handelt es sich um einen sogenannten Bakteraszites [45]. Dieser kann sowohl spontan reversibel sein als auch in eine SBP übergehen [29,102].

Zu den anerkannten Risikofaktoren für eine SBP gehören eine bereits durchgemachte SBP [46], ein geringer Eiweißgehalt im Aszites (< 1,5 g/dl) [100] sowie gastrointestinale Blutungen [55]. Ebenfalls begünstigend für eine SBP sind eine Erhöhung der Serumbilirubinkonzentration auf > 3,2 mg/dl und

eine Thrombopenie [50] sowie die Einnahme von Protonenpumpeninhibitoren [7]. Mit zunehmendem Model for End-stage Liver Disease (MELD)-Score steigt das Risiko einer SBP zusätzlich an [84]. Darüber hinaus konnte gezeigt werden, dass bestimmte genetische Polymorphismen im Toll-like receptor 2 Promotor und Intron 2 sowie bestimmte *NOD2* (nucleotide-binding oligomerisation domain 2)-Genotypen das Risiko einer SBP erhöhen [4,82].

2.1.2.3.2 *NOD2*-Gen

Der *NOD2*-Rezeptor ist ein Pattern Recognition Rezeptor, der vor allem von Leukozyten, aber auch von einigen Darmepithelzellen exprimiert wird.

NOD2-Risikovarianten werden hauptsächlich mit einer gestörten Schleimhautbarriere bei Morbus Crohn in Verbindung gebracht. Bei Trägern dieser *NOD2*-Varianten kann eine unzureichende Aktivierung von NF- κ B zu einer mangelhaften Eliminierung und verstärkten Translokation von Bakterien aus dem Darm führen [52,57].

Aber auch bei der Entwicklung von Komplikationen der Zirrhose scheinen *NOD2*-Varianten eine wichtige Rolle zu spielen: Studien konnten zeigen, dass bestimmte *NOD2*-Varianten (c.3020insC, p.R702W und p.G908R) Prädiktoren für die Entwicklung infektiöser Komplikationen und für ein erhöhtes Sterberisiko bei Patienten mit Zirrhose sind [4,21]. Mit den Varianten rs72796367 und rs5743271 konnten weitere Risikovarianten des *NOD2*-Gens identifiziert werden [56], jedoch ist deren Bedeutung in der Pathogenese hepatischer Erkrankungen noch unzureichend untersucht.

2.1.3 Prognose der Zirrhose und deren Abschätzung

Die Stadieneinteilung der Zirrhose nach D'Amico kann zur groben Abschätzung der Prognose genutzt werden. So haben kompensierte Patienten ohne Ösophagusvarizen (Stadium 1) eine 5-Jahres-Mortalität von etwa 1,5 %, mit Ösophagusvarizen (Stadium 2) steigt diese auf etwa 10 %. Nach dem ersten Dekompensationsereignis (ab Stadium 3) nimmt die 5-Jahres-Mortalität auf Werte > 20 % zu, ab dem zweiten Dekompensationsereignis auf ca. 88 % [33].

Mit dem Child-Pugh- und MELD-Score existieren zwei sehr häufig genutzte Scores, mit denen die Prognose von Patienten mit Zirrhose genauer eingeschätzt werden kann, allerdings sind sie nur für die kurzfristige Prognose der Mortalität aussagekräftig [10,27,28]. Der Child-Pugh-Score teilt die Zirrhose

anhand von Laborwerten (Gesamtbilirubin, Serumalbumin und International Normalized Ratio (INR) bzw. Quick-Wert) sowie dem Auftreten von Aszites und hepatischer Enzephalopathie in drei Stadien ein, die eine wichtige prognostische Aussagekraft haben (Tabelle 1) [91]. So haben Patienten im dritten Stadium eine 1-Jahres-Überlebenswahrscheinlichkeit von lediglich 40 %.

Tabelle 1: Child-Pugh Score

Punkte	1	2	3
Serumalbumin [g/dl]	> 3,5	2,8–3,5	< 2,8
Serumbilirubin [mg/dl]	< 2,0	2,0–3,0	> 3,0
Quick-Wert [%]	> 70	40–70	< 40
Aszites	kein	mäßig	viel
Hepatische Enzephalopathie	keine	Grad I–II	>Grad II
Child A: 5–6 Punkte; Child B: 7–9 Punkte; Child C: 10–15 Punkte			

Der MELD-Score [63,64] prognostiziert ebenfalls die Überlebenswahrscheinlichkeit von Patienten mit Lebererkrankungen im Endstadium und hat 2003 den Child-Pugh-Score bei der Einschätzung der Dringlichkeit einer Lebertransplantation abgelöst. Im Laufe der Zeit wurden mehrere Modifikationen des MELD-Scores vorgeschlagen, aber die Werte von Bilirubin, Kreatinin und INR gehen bei allen Varianten in die Berechnung ein. In der ursprünglichen Version wurde auch eine alkoholische Ätiologie der Erkrankung berücksichtigt. Da aber gezeigt werden konnte, dass diese weniger relevant für das Überleben ist, wurde die Ätiologie aus dem Modell entfernt.

Der MELD-Score berechnet sich wie folgt:

$$\text{MELD} = 10 \times \left(0,957 \times \ln \left(\text{Serumkreatinin} \left[\frac{\text{mg}}{\text{dl}} \right] \right) + 0,378 \times \ln \left(\text{Bilirubin gesamt} \left[\frac{\text{mg}}{\text{dl}} \right] \right) + 1,12 \times \ln(\text{INR}) + 0,643 \right)$$

In mehreren Studien konnte gezeigt werden, dass die Berücksichtigung des Serumnatriums die prädiktive Genauigkeit des MELD-Scores verbessert [13,69]. Daher wurde basierend auf dem MELD-Score der sogenannte MELD-Na-Score entwickelt, der wie folgt berechnet wird:

$$\text{MELD} - \text{Na} = \text{MELD} - \text{Serumnatrium} - (0.025 \times \text{MELD} \times (140 - \text{Serumnatrium})) \\ + 140$$

Der MELD-Plus-Score wurde ebenfalls als Verbesserung des MELD- bzw. MELD-Na-Scores vorgeschlagen. Er berücksichtigt neben den Parametern des MELD-Na-Scores Serumalbumin, Cholesterin, Leukozytenzahl, Alter sowie die Hospitalisierungsdauer [66].

Der Albumin-Bilirubin (ALBI)-Score kombiniert Serumalbumin- und Bilirubinwerte, um die Leberfunktion von Patienten sowohl mit hepatozellulärem Karzinom (HCC) [53,62,89,111] als auch ohne hepatozellulärem Karzinom [25,86,122] zu bewerten.

Der Fibrosis-4 (Fib-4) Index schätzt das Ausmaß einer Leberfibrose anhand von Alter, Alanin-Aminotransferase (ALAT, ALT), Aspartat-Aminotransferase (ASAT, AST) und Thrombozytenzahl ab [109]. Er wurde vielfach validiert und hat eine prognostische Aussagekraft für Patienten mit chronischen Lebererkrankungen.

Der CLIF Consortium Acute Decompensation (CLIF-C AD) [60] Score berechnet die Wahrscheinlichkeit, mit der hospitalisierte Patienten mit Zirrhose und akuter Dekompensation innerhalb des nächsten Jahres versterben. Die Berechnung erfolgt basierend auf Alter, Serumkreatinin, INR, Leukozytenzahl sowie Serumnatrium anhand folgender Formel:

$$\text{CLIF} - \text{C AD: } 10 \times (0.03 \times \text{Alter [Jahre]} + 0.66 \times \ln\left(\text{Kreatinin} \left[\frac{\text{mg}}{\text{dL}}\right]\right) + 1.71 \times \ln(\text{INR}) \\ + 0.88 \times \ln\left(\text{Leukozytenzahl} \left[\frac{10^9}{\text{l}}\right] - 0.05 \times \text{Na} \left[\frac{\text{mmol}}{\text{l}}\right] + 8\right))$$

Der 2022 veröffentlichte Early Prediction of Decompensation (EPOD)-Score ist der einzige Score, der sich nicht auf die Mortalität bezieht, sondern auf die Wahrscheinlichkeit einer Dekompensation. Dazu verwendet er die Thrombozytenzahl, sowie die Albumin- und Bilirubinkonzentration [104].

$$\text{EPOD} = \left((-0.55) \times \text{Albumin} \left[\frac{\text{g}}{\text{dl}}\right] + (-0.004) \times \text{Thrombozytenzahl} [10^3/\mu\text{L}] \right. \\ \left. + 0.16 \times \text{Bilirubin} \left[\frac{\text{mg}}{\text{dl}}\right] + 5.38 \right) \times 4$$

Bislang sind die Parameter, die eine Dekompensation prognostizieren, jedoch nicht vollständig definiert und es konnte sich noch keiner dieser Scores zur routinemäßigen Einschätzung einer Dekompensation durchsetzen.

2.2 Künstliche Intelligenz, Machine Learning, Deep Learning

Künstliche Intelligenz (KI, AI für englisch „Artificial Intelligence“) ist ein Teilgebiet der Informatik, das sich mit der Automatisierung von intelligentem Verhalten beschäftigt. Für den Begriff „künstliche Intelligenz“ existiert jedoch keine allgemeingültige Definition. Allgemein wird unter künstlicher Intelligenz der Versuch verstanden, menschliches Denken und Lernen durch Computer abzubilden, um diesen so Intelligenz zu verleihen.

Dabei kann zwischen starker und schwacher KI unterschieden werden. Unter einer starken KI versteht man Maschinen, die selbstständig Probleme unterschiedlicher Bereiche lösen können und damit das menschliche Verhalten in seiner Gesamtheit abbilden sollen. Teilweise werden auch Empathie und Bewusstsein als zentrale Eigenschaften einer starken KI betrachtet [87,107]. Schwache KI charakterisiert hingegen Algorithmen, die nur auf einen engen Themenbereich angewendet werden, um so bestimmte Probleme zu lösen [47]. Bislang wird nur schwache KI kommerziell oder in der Forschung genutzt, eine echte starke KI existiert (noch) nicht.

Maschinelles Lernen (Machine Learning, ML) ist ein wichtiges Teilgebiet der KI, das nicht nur in der Medizin zunehmend an Bedeutung gewinnt [103]. Beim ML lernen Computer aus Beispielen in einer Trainingsmenge ein statistisches Modell, welches anschließend verallgemeinert und auf neue Datensätze angewendet werden kann. Computer sollen so z. B. in bestehenden Datensätzen Muster/Gesetzmäßigkeiten erkennen oder Klassifikationsaufgaben bearbeiten [77,80]. Eng damit verwandt ist das Data Mining, bei dem aus vorgegebenen großen Datenmengen Muster extrahiert werden sollen.

Beim maschinellen Lernen kann man grundsätzlich zwischen überwachtem und unbeaufsichtigtem Lernen unterscheiden [103], wobei das überwachte Lernen deutlich häufiger genutzt wird. Hierbei werden dem Algorithmus in einer Trainingsmenge Ein- und Ausgabepaare zur Verfügung gestellt, anhand derer z. B. Entscheidungsfunktionen optimiert werden können. Diese Entscheidungsfunktion kann dann auf unbekannte Eingabedaten angewendet werden, um so einen vermuteten Ausgabewert zu berechnen. Abhängig vom gewählten Algorithmus können diese Verfahren sowohl zur Klassifikation als auch zur Regression genutzt werden. Bekannte Ansätze zum überwachten Lernen stellen Entscheidungsbäume (Decision Trees, DT), Random Forest Classifier (RFC), Support Vector Machines (SVM) und neuronale Netze dar.

Für das unüberwachte Lernen spielen u. a. Verfahren der Clusteranalyse mit dem Expectation-Maximization-Algorithmus eine wichtige Rolle.

Deep Learning wiederum bezeichnet eine mögliche Form von maschinellem Lernen. Beim Deep Learning werden v. a. künstliche neuronale Netze (KNN) mit Eingabe-, Ausgabeschicht und mehreren versteckten Schichten (hidden layers) genutzt, um maschinelles Lernen zu ermöglichen [6]. Im Vergleich zu klassischem maschinellem Lernen besitzen Deep Learning-Verfahren eine deutlich größere Anzahl freier Variablen, sodass sie zwar rechenintensiver und komplexer sind, aber auch wesentlich flexibler und daher besser geeignet für große, komplexe Datensätze (insbesondere in der Bild- oder Sprachverarbeitung) [72]. Durch verbesserte Algorithmen und Hardware hat Deep Learning in den letzten 10 Jahren deutlich an Popularität gewonnen.

Maschinelles Lernen bzw. Deep Learning kann in der Medizin auf zwei Arten eingesetzt werden: Zum einen kann ML zur automatisierten Auswertung bzw. Interpretation von Daten genutzt werden, die normalerweise von Menschen durchgeführt wird. Zum anderen kann es aber auch zur Extraktion von Merkmalen/Mustern aus komplexen Datenmengen verwendet werden, welche für Menschen sonst nur schwierig zu erkennen wären [97].

Obwohl maschinelles Lernen in allen medizinischen Fachgebieten insbesondere in der Forschung zunehmend an Bedeutung gewinnt, ist die tatsächliche klinische Relevanz noch gering. Dafür existiert vermutlich eine Reihe von Gründen, z. B. fehlt in den allermeisten Fällen eine externe und prospektive Validierung der Modelle und die Datensätze sind oft sehr heterogen und wenig/nicht standardisiert. Bei vielen Mitarbeitern im Gesundheitswesen ist das Wissen über maschinelles Lernen begrenzt und daher fehlt häufig das Vertrauen in die Verfahren. Auch treten immer wieder Datenschutzbedenken auf [81].

2.3 Künstliche Intelligenz und maschinelles Lernen in der Hepatologie

Seit 2018 ist ein exponentieller Anstieg an Studien, die den Einsatz von maschinellem Lernen in der Hepatologie untersuchen, zu verzeichnen [81].

Im Bereich der Hepatologie wurde maschinelles Lernen bisher v. a. im Bereich der Diagnostik getestet, z. B. zur Befundung bildgebender Verfahren oder histologischer Proben. Am häufigsten wurde maschinelles Lernen zur automatisierten Diagnose von Lebererkrankungen in Bilddaten eingesetzt, seltener zur Einschätzung der Prognose. So existieren einige Studien, in denen die Detektion

hepatozellulärer Karzinome in Computertomographie (CT)- bzw. Magnetresonanztomographie (MRT)-Aufnahmen mittels maschinellen Lernens untersucht wird. Zudem wird maschinelles Lernen eingesetzt, um das Ausmaß einer Fibrose in verschiedenen Untersuchungen abzuschätzen [26,37,108,119]. Ein Beispiel hierfür bietet die Studie von Bosch et al., bei der histologische Proben der Leber mithilfe eines neuronalen Netzes analysiert wurden, um den hepatisch-venösen Druckgradienten (HVPG) zu bestimmen [16]. Bezüglich der Prognoseeinschätzung wurden beispielsweise Modelle zur Vorhersage des Rezidivrisikos hepato- oder cholangiozellulärer Karzinome [61,92] und zur Einschätzung der Mortalität von Lebertransplantationskandidaten [11] bzw. zur Prognose eines Transplantatversagens entwickelt [74]. Außerdem gibt es mehrere Studien, die bei Patienten mit chronischer Hepatitis-B- oder -C-Virusinfektion das Fortschreiten der Krankheit und die Ausbildung einer Fibrose untersuchen [71,113,117].

Zur Vorhersage der Mortalität bzw. Bestimmung der Dekompensationswahrscheinlichkeit bei Patienten mit Zirrhose existieren gegenwärtig jedoch weniger Studien [8,30,76,120]. Beispielsweise testeten Guo et al. verschiedene ML-Techniken zur Schätzung der 90-, 180- und 360-Tage-Mortalität von Patienten mit Leberzirrhose [51].

2.4 Ziel der Studie

Obwohl maschinelles Lernen in der Hepatologie zunehmend auch für prognostische Zwecke eingesetzt wird, wurde noch keine Studie publiziert, die den Einsatz von maschinellern Lernen zur Vorhersage von Dekompensationen testet. Außerdem wurden genetische Daten wie *NOD2*-Genotypen in den bisherigen Studien nur selten betrachtet.

Da die Parameter, die auf eine Dekompensation hinweisen, noch nicht vollständig definiert sind, soll in dieser Arbeit untersucht werden, ob mithilfe von maschinellern Lernen Parameter identifiziert werden können, die mit Dekompensationen assoziiert sind.

3 Patienten und Methodik

3.1 Patienten

Als Grundlage für die Studie wurde die Datenbank der INCA-Studie (Impact of *NOD2* genotype-guided antibiotic prevention on survival in patients with liver Cirrhosis and Ascites) genutzt. Im Rahmen der INCA-Studie wird untersucht, ob das Überleben von Patienten mit Zirrhose und *NOD2*-Varianten als genetisch definierte Hochrisikogruppe durch eine Primärprophylaxe einer SBP mit Norfloxacin verbessert werden kann [23]. Ursprünglich enthielt die Datenbank Daten von 1.415 Patienten aus drei deutschen Universitätskliniken (Homburg, Halle und Jena). Da alle Patienten aus Jena bereits initial dekompensiert waren und viele fehlende Daten aufwiesen, wurden sie aus der Analyse ausgeschlossen. Die verbleibende Datenbank umfasst 983 Patienten aus Homburg und Halle und besteht sowohl aus einem prospektiven als auch einem retrospektiven Teil.

Für den prospektiven Teil der Datenbank wurden die Patienten während eines medianen Follow-up von 12 Monaten und eines maximalen Follow-up von 55 Monaten beobachtet. Daran nahmen 825 Patienten (83,9 %) teil.

Der retrospektive Teil enthält anamnestische und klinische Daten aus Klinikinformationssystemen, sowie insbesondere auch Angaben über vorhergehende Dekompensationen. Zum Inklusionszeitpunkt wurden die aktuellen Laborwerte der Studienteilnehmer erhoben; für die Patienten aus Homburg wurden zudem die minimalen und maximalen bekannten Laborwerte aus dem Klinikinformationssystem ermittelt.

Die Diagnose einer Zirrhose wurde anhand einer (i) Biopsie, (ii) Kombination von klinischen, laborchemischen, sonographischen und endoskopischen Untersuchungen oder (iii) Lebersteifigkeit > 13.0 kPa gestellt. Bei einer Lebersteifigkeit < 19.7 kPa wurde die Diagnose zusätzlich durch (i) oder (ii) bestätigt [23,95]. Patienten mit schweren Vorerkrankungen, wie z. B. einer chronischen Herzinsuffizienz (Ejektionsfraktion < 20 %) oder nicht-resektablen Tumoren mit Ausnahme des hepatozellulären Karzinoms, wurden ausgeschlossen. In dem Kollektiv war die häufigste Ätiologie Alkoholismus (54 %), gefolgt von einer chronischen Hepatitis C-Virus-Infektion (14 %).

Dekompensation wurde definiert durch das Auftreten von Aszites, Ikterus, hepatischer Enzephalopathie oder einer Varizenblutung, Rekompensation durch die Abwesenheit des definierenden Dekompensationskriteriums. Aszites wurde dabei anhand einer körperlichen Untersuchung oder Sonografie diagnostiziert, die hepatische Enzephalopathie nach den West-Haven-Kriterien [39] und Ikterus durch eine Serumbilirubinkonzentration ≥ 3 mg/dl oder einen Bilirubinanstieg um ≥ 3 mg/dl im

Vergleich zur Voruntersuchung. Die Diagnose einer Varizenblutung wurde gemäß den Baveno-Kriterien [35] anhand von Hämatemesis/Meläna oder mithilfe einer Endoskopie gestellt [23,95].

Fehlende kategoriale Werte wurden bei allen Patienten durch -1 ersetzt. Fehlende numerische Werte wurden bei den Homburger Patienten durch entsprechende Daten aus dem Klinikinformationssystem (i.s.h.med, SAP) ersetzt, für die Patienten aus Halle wurde jeweils der Mittelwert der Werte der übrigen Patienten benutzt.

Tabelle 2 zeigt die Charakteristika der kategorialen Variablen des Datensets, Tabelle 2: Charakteristika der kategorialen Variablen

die der numerischen Variablen. Die Angabe „Min“ (Minimum) bzw. „Max“ (Maximum) hinter den Variablennamen gibt an, dass der jeweils niedrigste bzw. höchste Wert des Laborparameters bis zum Zeitpunkt der Aufnahme in die Studie berücksichtigt wurde.

Tabelle 2: Charakteristika der kategorialen Variablen

Parameter		Total (N = 983)	Homburg (N = 771)	Halle (N = 212)
Frühere Dekompensation	1 (zum Indexzeitpunkt kompensiert, keine frühere Dekompensation)	313 (31,8 %)	267 (35 %)	46 (22 %)
	2 (zum Indexzeitpunkt kompensiert, aber frühere Dekompensation)	354 (36,0 %)	188 (24 %)	166 (78 %)
	3 (zum Indexzeitpunkt dekompenziert)	316 (32,1 %)	316 (41 %)	0 (0 %)
Dekompensation während Follow-up	-1 (unbekannt)	158 (16,1 %)	151 (20 %)	7 (3 %)
	0 (nein)	441 (44,9 %)	392 (51 %)	49 (23 %)
	1 (ja)	384 (39,1 %)	228 (30 %)	156 (74 %)
Geschlecht	0 (männlich)	647 (65,8 %)	510 (66 %)	137 (65 %)
	1 (weiblich)	336 (34,2 %)	261 (34 %)	75 (35 %)
Ätiologie	1 (Alkoholismus)	531 (54,0 %)	370 (48 %)	161 (76 %)
	2 (HCV)	134 (13,6 %)	130 (17 %)	4 (2 %)
	3 (HBV)	29 (3,0 %)	28 (4 %)	1 (0 %)
	4 (nichtalkoholische Fettlebererkrankung, NAFLD)	75 (7,6 %)	57 (7 %)	18 (8 %)
	5 (Autoimmunhepatitis)	31 (3,2 %)	20 (3 %)	11 (5 %)
	6 (PBC)	10 (1,0 %)	7 (1 %)	3 (1 %)
	7 (PSC)	17 (1,7 %)	15 (2 %)	2 (1 %)
	8 (Hämochromatose)	12 (1,2 %)	12 (2 %)	0 (0 %)
	9 (M. Wilson)	2 (0,2 %)	2 (0 %)	0 (0 %)
	10 (kryptogen)	101 (10,3 %)	91 (12 %)	10 (5 %)
	11 (andere)	41 (4,2 %)	38 (5 %)	2 (1 %)
Akut-auf-chronisches Leberversagen (ACLF) zu Beginn	-1 (unbekannt)	1 (0,1 %)	0 (0 %)	1 (0 %)
	0 (kein ACLF)	898 (91,4 %)	712 (92 %)	186 (88 %)
	1 (ACLF 1)	66 (6,7 %)	50 (6 %)	16 (8 %)
	2 (ACLF 2)	15 (1,5 %)	7 (1 %)	8 (4 %)
	3 (ACLF 3)	3 (0,3 %)	2 (0 %)	1 (0 %)
Diabetes mellitus	0 (nein)	668 (68,0 %)	532 (69 %)	136 (64 %)
	1 (ja)	315 (32,0 %)	239 (31 %)	76 (36 %)

rs5743271	-1 (unbekannt)	226 (23,0 %)	14 (2 %)	212 (100 %)
	0 (nein)	742 (75,5 %)	742 (96 %)	0 (0 %)
	1 (heterozygot)	115 (11,7 %)	15 (2 %)	0 (0 %)
rs72796367	-1 (unbekannt)	227 (23,1 %)	15 (2 %)	212 (100 %)
	0 (nein)	720 (73,2 %)	720 (93 %)	0 (0 %)
	1 (heterozygot)	35 (3,6 %)	35 (5 %)	0 (0 %)
	2 (homozygot)	1 (0,1 %)	1 (0 %)	0 (0 %)
NOD2 p.R702W	-1 (unbekannt)	3 (0,3 %)	3 (0 %)	182 (86 %)
	0 (nein)	862 (87,7 %)	680 (88 %)	29 (14 %)
	1 (heterozygot)	117 (11,9 %)	88 (11 %)	1 (0 %)
	2 (homozygot)	1 (0,1 %)	3 (0 %)	206 (97 %)
NOD2 p.G908R	-1 (unbekannt)	3 (0,3 %)	743 (96 %)	6 (3 %)
	0 (nein)	949 (96,5 %)	24 (3 %)	0 (0 %)
	1 (heterozygot)	30 (3,1 %)	1 (0 %)	0 (0 %)
	2 (homozygot)	1 (0,1 %)	3 (0 %)	195 (92 %)
NOD2 c.3020insC	-1 (unbekannt)	3 (0,3 %)	728 (94 %)	16 (8 %)
	0 (nein)	923 (93,9 %)	40 (5 %)	1 (0 %)
	1 (heterozygot)	56 (5,7 %)	0 (0 %)	0 (0 %)
	2 (homozygot)	1 (0,1 %)	14 (2 %)	212 (100 %)
Varizen	-1 (unbekannt)	9 (0,9 %)	0 (0 %)	9 (4 %)
	0 (nein)	472 (48,0 %)	408 (53 %)	139 (66 %)
	1 (ja)	502 (51,1 %)	363 (47 %)	64 (30 %)
Protonenpumpen- inhibitoren (PPI)	-1 (unbekannt)	11 (1,1 %)	9 (4 %)	8 (4 %)
	0 (nein)	290 (29,5 %)	139 (66 %)	56 (26 %)
	1 (ja)	682 (69,4 %)	64 (30 %)	148 (70 %)
Betablocker	-1 (unbekannt)	11 (1,1 %)	3 (0 %)	8 (4 %)
	0 (nein)	473 (48,1 %)	419 (54 %)	124 (58 %)
	1 (ja)	499 (50,8 %)	349 (45 %)	80 (38 %)
Langfristige Antibiotikaeinnahme	-1 (unbekannt)	12 (1,2 %)	5 (1 %)	7 (3 %)
	0 (nein)	797 (81,1 %)	625 (81 %)	172 (81 %)
	1 (ja)	174 (17,7 %)	141 (18 %)	33 (16 %)

Tabelle 3: Charakteristika der numerischen Variablen

Parameter	Mittelwert	Standardabweichung
Alpha-Fetoprotein (AFP) Max (IU/ml)	1344,7	10050,5
Albumin (g/l)	35,34	7,27
Albumin Min (g/l)	27,08	7,91
Alkalische Phosphatase (AP) Max (U/l)	255,33	209,28
ALT (U/l)	63,27	148,6
ALT Max (U/l)	278,78	563,14
Ammoniak Max (µg/dl)	136,79	78,93
AST (U/l)	94,25	155,54
AST Max (U/l)	564,18	2094,27
Bilirubin (mg/dl)	2,47	4,21
Bilirubin Max (mg/dl)	5,84	6,82
Calcium (Ca) (mmol/l)	2,27	0,18
Calcium Max (mmol/l)	2,49	0,18
Calcium Min (mmol/l)	2	0,2
Cholesterol Max (mg/dl)	195,92	58,4
Cholinesterase (CHE) Min (kU/l)	3,34	2,25
C-reaktives Protein (CRP) (mg/l)	19,56	31,26
CRP Max (mg/l)	94,73	86,2
Ferritin Max (ng/ml)	1099,24	2576,17
Ferritin Min (ng/ml)	320,74	658,6
Glucose Max (mg/dl)	222,74	115,32
Harnstoff (mg/dl)	43,58	31,47
Harnstoff Max (mg/dl)	92,89	58,94
Hämoglobin (HB) (g/dl)	12,13	2,54
HB Min (g/dl)	9,3	2,53
INR	1,29	0,39
INR Max	1,77	0,72
Kalium (K) (mmol/l)	4,12	0,54
Kalium Max (mmol/l)	5,12	0,52
Kalium Min (mmol/l)	3,38	0,42
Kreatinin (mg/dl)	1,13	0,75
Kreatinin Max (mg/dl)	2,12	1,51
Leukozyten Min (x10 ⁹ /l)	4,12	1,9
Leukozytenzahl (x10 ⁹ /l)	7,07	4,01

Natrium (Na) (mmol/l)	137,7	4,7
Natrium Min (mmol/l)	131,7	4,3
Partielle Thromboplastinzeit (PTT) (s)	31,78	7,8
Partielle Thromboplastinzeit Max (s)	52,95	27,23
Thrombozyten Min ($\times 10^9/l$)	88,46	52,99
Thrombozytenzahl ($\times 10^9/l$)	154,31	102,6
Triacylglyceride (TAG) Max (mg/dl)	193,99	172,39
γ -Glutamyltransferase (GGT) Max (U/l)	544,65	601,35

3.2 Methodik

3.2.1 Grundlagen

3.2.1.1 Machine Learning-Verfahren

Zur Bestimmung der Parameter, die eine Dekompensation prognostizieren, wurden verschiedene maschinelle Lernverfahren getestet: Entscheidungsbäume, Random Forest Classifier, Support Vector Machines und neuronale Netze.

3.2.1.1.1 Entscheidungsbäume

Entscheidungsbäume sind gerichtete Bäume, die hierarchisch angeordnete Entscheidungsregeln in einer übersichtlichen Struktur darstellen, um Objekte zu klassifizieren [19]. Objekte werden dabei - wie im ML üblich - durch eine Menge von Attribut-Wert-Paaren beschrieben. Ein Entscheidungsbaum enthält genau einen Wurzelknoten, beliebig viele innere Knoten und mindestens zwei Blätter. Jeder innere Knoten steht dabei stellvertretend für eine logische Entscheidungsregel, die Blattknoten repräsentieren die möglichen Lösungen des Problems bzw. Ausprägungen der Zielvariablen.

In jedem inneren Knoten wird ein bestimmtes Attribut a betrachtet und dessen Wert für ein zu klassifizierendes Objekt abgefragt. Die von dem jeweiligen Knoten ausgehenden Kanten sind mit den möglichen Werten von a markiert. Um ein unbekanntes Objekt zu klassifizieren, wird ein passender Weg (Pfad) durch den Baum von der Wurzel bis zu einem Blattknoten gesucht. Dazu wird an jedem inneren Knoten die Ausprägung des betrachteten Attributs a abgefragt und die entsprechende Kante zum nächsten Knoten gewählt. Wenn ein Blattknoten erreicht wird, wird dessen zugehöriger Wert als Ergebnis der Klassifikation zurückgegeben.

Zum Aufbau von Entscheidungsbäumen existieren zahlreiche Algorithmen, CART [19], ID3 [93] und sein Nachfolger CD 4.5 [94] (bzw. CD 5.0) sind jedoch am populärsten. Den Kern aller Algorithmen bildet die Auswahl des jeweils nächsten Attributs. Da der Entscheidungsbaum möglichst klein gehalten werden soll, wäre ein ideales Attribut eines, das die verbleibenden Beispiele genau in eine positive und negative Beispielmenge aufteilt.

Zur Attributauswahl wird meistens die Entropie bzw. der davon abgeleitete Informationsgewinn (Information Gain, IG) oder alternativ der Gini-Koeffizient verwendet.

ID3 [93] steht für „Iterative Dichotomiser 3“ und beschreibt ein Verfahren zur Konstruktion von Entscheidungsbäumen für relativ große Datenmengen mit einer hohen Anzahl möglicher Attribute. In jedem Schritt wird das Attribut mit der niedrigsten Entropie bzw. dem höchsten Informationsgewinn bestimmt. Aus diesem Attribut wird eine neue Entscheidungsregel konstruiert und dementsprechend ein neuer (innerer) Knoten zum Entscheidungsbaum hinzugefügt. Wenn alle Elemente der Trainingsmenge klassifiziert wurden, terminiert der Algorithmus.

C4.5 [94] ist eine Weiterentwicklung des ID3-Algorithmus, bei dem anstatt des IG bzw. der Entropie ein normierter IG betrachtet wird.

Der entscheidende Vorteil von Entscheidungsbäumen liegt in ihrer Einfachheit, Übersichtlichkeit und Erklärbarkeit, der größte Nachteil besteht in ihrer sehr hohen Anfälligkeit für Overfitting, da sie sich sehr stark an die Daten der Trainingsmenge anpassen.

3.2.1.1.2 Random Forests

Random Forests [20,54] sind Mengen von möglichst unkorrelierten Entscheidungsbäumen, die zur Klassifikation und Regression eingesetzt werden können. Während des Trainingsprozesses wird nicht mehr ein einzelner optimaler Entscheidungsbaum generiert, sondern mithilfe von verschiedenen Randomisierungsverfahren wird eine Menge von Bäumen erstellt, die in einem Ensemble zur Entscheidungsfindung genutzt werden. Die Ergebnisse aller Entscheidungsbäume werden kombiniert

und die Klasse mit den meisten „Stimmen“ bzw. mit der höchsten Gesamtwahrscheinlichkeit wird ausgewählt.

Zur Randomisierung können unterschiedliche Techniken verwendet werden, meist wird eine Kombination aus Bagging und Feature Randomness eingesetzt. Bei der Feature Randomness berücksichtigt jeder Baum an den Splits, also den inneren Knoten, nur eine zufällige Teilmenge der Variablen und wählt daraus den besten Parameter aus. Dies führt zu einer geringeren Korrelation zwischen den Bäumen und damit zu einer höheren Diversität und besseren Verallgemeinerbarkeit. Bagging steht für Bootstrap Aggregation. Dabei werden aus den Trainingsdatensatz zunächst n Bootstrap samples generiert, auf jedem dieser Bootstrap samples wird ein eigener Baum trainiert [75]. Wie auch bei der Feature Randomness kann hierbei aus der Variablenmenge zunächst eine zufällige Teilmenge ausgewählt werden, aus welcher dann die beste Variable für den Knoten ausgewählt wird. Zur Bestimmung der Vorhersage werden die Teilergebnisse aller n Bäume betrachtet, das Mehrheitsvotum entscheidet.

Durch die Bildung einer Menge unkorrelierter Bäume sollen einige Nachteile von klassischen Entscheidungsbäumen ausgeglichen werden: Einzelne Entscheidungsbäume passen sich sehr stark an die Daten der Trainingsmenge an und tendieren daher enorm zu Overfitting. Bereits kleine Veränderungen der Trainingsmenge lösen so große Veränderungen in einem Entscheidungsbaum aus, daher ist ihre Verallgemeinerbarkeit gering. Durch die zufällige Varianz in den einzelnen Entscheidungsbäumen soll das Overfitting deutlich reduziert werden, was zu einer höheren Genauigkeit insbesondere bei unbekannten Daten führt.

Da sich die einzelnen Bäume parallel aufbauen bzw. auswerten lassen, sind die Trainings- und Evaluierungszeiten vergleichsweise kurz. Daher sind Random Forests auch für große Datenmengen sehr effizient.

Im Gegensatz zu neuronalen Netzen sind die Entscheidungen von Random Forests gut erklärbar bzw. nachvollziehbar, und es lassen sich auch Aussagen über Zusammenhänge zwischen Klassen treffen.

3.2.1.1.3 Support Vector Machines

Support Vector Machines können sowohl zur Klassifikation als auch zur Regression eingesetzt werden. Dabei handelt es sich um einen Large Margin Classifier, d. h. die Objekte sollen so in Klassen aufgeteilt werden, dass ein möglichst breiter Rand um die Klassen herum frei von Objekten bleibt.

Support Vector Machines arbeiten auf Vektoren in Vektorräumen: Jedes Objekt aus der Trainings-/ Testmenge wird durch einen Vektor repräsentiert. Durch die Support Vector Machine soll nun eine

Hyperebene gefunden werden, die die Objekte in unterschiedliche Klassen teilt [41]. Dabei werden nur die Vektoren, die der Hyperebene jeweils am nächsten liegen (Stützvektoren, Support Vectors), berücksichtigt, alle anderen Vektoren (Objekte) beeinflussen Form und Lage der Trennebene nicht. Ziel ist es, den Abstand der Stützvektoren von der Hyperebene zu maximieren, damit ein möglichst breiter, leerer Rand entsteht, um auch unbekannte Objekte zuverlässig klassifizieren zu können [115].

Problematisch ist bei diesem Vorgehen, dass eine saubere Trennung der Klassen durch eine Hyperebene nur bei linear trennbaren Objekten möglich ist; diese Voraussetzung ist in realen Anwendungen jedoch i. d. R. nicht erfüllt. Abhilfe schafft in diesem Fall der sogenannte Kerneltrick [17]: Die grundsätzliche Idee besteht darin, den Vektorraum mit den Trainingsobjekten in höherdimensionale Räume zu überführen, da dort auch eigentlich nicht linear trennbare Mengen linear trennbar werden [105]. Im passenden höherdimensionalen Raum wird wiederum eine Hyperebene bestimmt, diese wird bei der Rücktransformation in den ursprünglichen niedrigerdimensionalen Raum zu einer nichtlinearen Hyperfläche, die die Vektoren in verschiedene Klassen unterteilt. Da der Prozess der Transformation allerdings sehr rechenintensiv und die Darstellung der Hyperfläche im niedrigdimensionalen Raum extrem komplex ist, ist dieses Vorgehen praktisch nicht nutzbar. Statt der tatsächlichen Berechnung der Transformationen werden zur Beschreibung der Trennfläche geeignete Kernelfunktionen genutzt, die in höherdimensionalen Räumen die entsprechenden Hyperebenen beschreiben, aber im ursprünglichen Vektorraum „nutzbar“ bleiben.

Durch Kernelparameter lassen sich die Flexibilität und Komplexität des Modells gezielt steuern: Der Parameter C wirkt als Regularisierungsparameter, der einen Kompromiss zwischen der maximalen Trennung der Klassen und der Reduzierung von Fehlklassifikationen im Training herstellt. Der Parameter γ legt fest, wie stark einzelne Datenpunkte die Form der Entscheidungsgrenzen beeinflussen, also wie stark das Modell insgesamt durch einzelne Trainingsbeispiele geprägt wird. Durch eine Feinjustierung dieser Kernelparameter soll eine optimale Balance zwischen Genauigkeit und Generalisierbarkeit erreicht werden.

3.2.1.1.4 Neuronale Netze

Neuronale Netze gehören momentan zu den populärsten maschinellen Lernverfahren, da sie sehr gut auf großen Mengen unstrukturierter Daten arbeiten. Grundsätzlich lernt bzw. optimiert ein neuronales Netz eine Funktion $f(\cdot): R^m \rightarrow R^o$ zwischen der Menge der m Eingabeparameter und der Menge der o Ausgabeparameter [6]. Es besteht aus einer Menge von Knoten, den Neuronen, die in mehreren Schichten angeordnet sind. Obwohl der genaue Aufbau neuronaler Netze abhängig vom jeweiligen Einsatzzweck erheblich variieren kann, existiert stets eine Eingabeschicht mit den Neuronen

$\{x_{1i} | x_{11}, \dots, x_{1m}\}$, die die m Parameter des Datensatzes repräsentieren, und eine Ausgabeschicht, die das Ergebnis liefert. Dazwischen können beliebig viele „versteckte“ Schichten (hidden layers) liegen. In den versteckten Schichten nimmt jedes Neuron die Werte $x_1, \dots, x_m$ aus der vorhergehenden Schicht entgegen und berechnet mithilfe einer Propagierungsfunktion einen Nettoeingabewert, im einfachsten Fall wird dafür eine gewichtete Summe $net_j = \sum_{i=1}^n w_{ij} \cdot o_i$ benutzt [6]. Die benötigten Gewichte w_{ij} werden nicht vorgegeben, sondern im Laufe des Trainingsprozesses gelernt und angepasst.

Der berechnete Nettoeingabewert net_j geht anschließend in eine Aktivierungsfunktion ein, die bestimmt, ob bzw. wie stark das Neuron aktiviert wird. Die einfachste mögliche Aktivierungsfunktion stellt eine Schwellenwertfunktion $a_j(net_j) = \begin{cases} 1, & \text{wenn } net_j > \theta_j \\ 0, & \text{sonst} \end{cases}$ dar, die in Abhängigkeit vom Schwellenwert θ_j nur zwischen zwei Zuständen (1 = aktiviert, 0 = nicht aktiviert) unterscheidet [6,121]. Nachteil der einfachen Schwellenwertfunktion (auch Standard-Binärfunktion genannt) ist die fehlende Stetigkeit und damit auch fehlende Differenzierbarkeit, welche Voraussetzung für den Backpropagation-Algorithmus ist. Stattdessen werden in der Praxis häufig sigmoide Funktionen wie die logistische Aktivierungsfunktion oder Tangenshyperbolicusfunktion eingesetzt [3]. Dabei ist zu beachten, dass die verwendeten Aktivierungsfunktionen nichtlinear sind, da das neuronale Netz andernfalls auf Ein- und Ausgabeschicht reduziert werden könnte [121]. Die Nichtlinearität von neuronalen Netzen kommt nämlich durch die versteckten Schichten und insbesondere durch den Einsatz nichtlinearer Aktivierungsfunktionen zustande.

Analog zu den Gewichten der Propagierungsfunktion können die Parameter der Aktivierungsfunktion (z. B. θ_j im Fall der einfachen Schwellenwertfunktion) im Laufe des Lernprozesses optimiert werden. Für den Trainingsprozess wird außerdem noch eine Fehlerfunktion E zur Berechnung der Abweichung zwischen ermitteltem Ausgabewert o_i und Sollausgabewert t_i benötigt. Hierfür kann die quadratische Fehlerfunktion $E = \sum_{i=1}^a (o_i - t_i)^2$ verwendet werden, wobei a die Anzahl der Neuronen in der Ausgabeschicht angibt. Durch die Quadrierung wird verhindert, dass sich Abweichungen vom Sollwert mit gegensätzlichen Vorzeichen aufheben.

Der grundsätzliche Lernprozess in einem vorwärtsgerichteten neuronalen Netz basiert auf dem Backpropagation-Algorithmus.

Vereinfacht kann der Backpropagation-Algorithmus wie folgt beschrieben werden [6,101]:

1. Zufällige Initialisierung der Startgewichte

Die Gewichte w_{ij} zwischen den Neuronen werden mit zufälligen Werten initialisiert.

2. Berechnung der Ausgabe

Für jedes Trainingsbeispiel p wird die Ausgabe des neuronalen Netzes berechnet. Dafür werden u. a. die Propagierungs- und Aktivierungsfunktion benutzt.

3. Berechnung des Fehlers

Mithilfe der gewählten Fehlerfunktion E wird berechnet, wie stark die Ausgabe des Netzes basierend auf den aktuellen Gewichten w_{ij} von der erwarteten Ausgabe abweicht. Ziel des Backpropagation-Algorithmus ist es, den durchschnittlichen Gesamtfehler über alle Trainingsbeispiele durch eine Modifikation der Gewichte w_{ij} im Laufe des Trainingsprozesses zu minimieren, indem die Gewichte einzeln für jeden Datensatz geändert werden.

4. Bestimmung der Suchrichtung

Für die Minimierung des Gesamtfehlers des Netzes wird die Fehlerfunktion E als Zielfunktion eines Optimierungsproblems betrachtet. Mithilfe des Gradientenverfahrens wird nun für jedes Gewicht w_{ij} berechnet, in welche Richtung es verändert werden muss, um den Gesamtfehler am stärksten zu minimieren. Der Gradient der Fehlerfunktion zeigt stets in Richtung der aktuell steilsten Stelle. Ausgehend von der Ausgabeschicht werden alle Fehlersignale rekursiv berechnet, sodass die Beiträge der vorgelagerten Gewichte ermittelt werden können.

5. Bestimmung der Schrittweite/Lernrate η

Nachdem im letzten Schritt bestimmt wurde, in welche Richtung die Gewichte verändert werden müssen, wird nun mithilfe der Schrittweite bzw. Lernrate η bestimmt, wie stark die jeweiligen Gewichte geändert werden. Die Lernrate hat einen großen Einfluss auf die Konvergenz und Performance des Algorithmus. Bei einer zu kleinen Lernrate werden die Gewichte immer nur gering verändert, weshalb viele Iterationen benötigt werden, sodass der Zeitaufwand enorm steigt. Bei einer zu großen Lernrate hingegen besteht die Gefahr, dass ein (lokales) Minimum der Fehlerfunktion evtl. übersprungen wird. Für die genaue Wahl der Lernrate existieren keine allgemeingültigen Empfehlungen, sondern sie muss üblicherweise ausprobiert werden.

6. Änderung der Gewichte:

Die Gewichte werden wie folgt angepasst: $\Delta w_{ij} = \eta \cdot \delta_{pj} \cdot net_{pi}$

7. Berechnung der neuen Ausgabewerte

Analog zum zweiten Schritt wird die Ausgabe des neuronalen Netzes basierend auf den neuen Gewichten w'_{ij} berechnet.

8. Berechnung des neuen Gesamtfehlers

Wie im dritten Schritt wird der neue Gesamtfehler berechnet.

9. Überprüfung der Abbruchkriterien

Sofern das Abbruchkriterium erfüllt ist, terminiert der Backtracking-Algorithmus, andernfalls wird wieder mit Schritt 4 fortgefahren.

Grundsätzlich sind unterschiedliche Abbruchkriterien denkbar, z. B. falls die Fehlerfunktion ein lokales Minimum erreicht oder falls die absolute Veränderung der Fehlerfunktion einen bestimmten Wert unterschreitet.

3.2.1.2 Inspektion des Modells: Permutation Feature Importance

Viele Machine Learning-Verfahren werden als Black Boxes betrachtet, die aus einer Eingabe eine Ausgabe generieren, ohne dass man nachvollziehen kann, warum das Verfahren gerade zu diesem Ergebnis gekommen ist. Welche Parameter wie stark gewichtet werden und wie diese sich gegenseitig beeinflussen, bleibt dabei im Allgemeinen unbekannt.

Allerdings ist es gerade in der Medizin wichtig zu wissen, wie eine Entscheidung zustande gekommen ist und wie stark die einzelnen Parameter gewichtet werden, um einschätzen zu können, ob eine Entscheidung bzw. das zugrundeliegende Modell plausibel sind. Daher wurden verschiedene Modellinspektionstechniken entwickelt, die einen Einblick in die genaue Arbeitsweise des Modells zulassen und Rückschlüsse auf die verwendeten Variablen erlauben sollen.

Die Permutation Feature Importance (PFI) ist eine Technik zur Erklärung von Klassifizierungs- und Regressionsmodellen, die von Leo Breimann entwickelt wurde [20]. Dabei wird untersucht, wie stark sich der Score eines Modells (z. B. die Genauigkeit) ändert, wenn der Wert einer bestimmten Variablen zufällig verändert wird. Je stärker der Score abfällt, desto wichtiger ist der Parameter für das Modell.

Das Hauptproblem der PFI ist, dass sie sehr stark durch korrelierte Parameter verzerrt wird. Der Einfluss korrelierter Parameter auf die Berechnung der PFI kann jedoch reduziert werden, indem die Parameter mithilfe einer Clusteranalyse in verschiedene Cluster aufgeteilt werden und pro Cluster jeweils nur ein Parameter für die Analyse verwendet wird.

In der vorliegenden Arbeit wurde mit der Ward-Methode ein hierarchisches agglomeratives Clusterverfahren verwendet. Bei den agglomerativen Verfahren bildet zunächst jedes Objekt einen eigenständigen Cluster, diese werden schrittweise nach bestimmten Kriterien (im Allgemeinen Abstands- oder Ähnlichkeitsmaße) fusioniert. Bei der Ward-Methode werden die Cluster immer so verschmolzen, dass der Zuwachs der totalen Varianz minimal ist [116].

3.2.1.3 Metriken zur Evaluation der Verfahren

Um die Güte der Verfahren beurteilen zu können, existieren unterschiedliche Metriken, u. a. die Accuracy, Precision und Recall bzw. davon abgeleitet der F1-Score sowie die Area under the Curve (AUC) der Receiver Operating Characteristic (ROC)-Kurve [90].

Die Accuracy (Genauigkeit) ist die am häufigsten benutzten Metrik. Sie ist definiert als Anteil der korrekten Vorhersagen, entspricht also dem Quotienten aus der Anzahl korrekter Vorhersagen und der Gesamtzahl der Vorhersagen.

$$accuracy(y, \hat{y}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} 1(\hat{y}_i = y_i)$$

Die Precision (Präzision) entspricht dem positiven prädiktiven Wert, beschreibt also die Fähigkeit des Klassifikators, einen negativen Testfall nicht als positiv zu beschreiben bzw. wie viele der positiven Vorhersagen tatsächlich korrekt sind. Sie berechnet sich wie folgt:

$$precision = \frac{true\ positive}{true\ positive + false\ positive}$$

Der Recall (Trefferquote) entspricht der Sensitivität, gibt also die Wahrscheinlichkeit an, mit der ein positives Beispiel auch als solches erkannt wird.

$$recall = \frac{true\ positive}{true\ positive + false\ negative}$$

Der F1-Score berechnet das gewichtete harmonische Mittel von Precision und Recall:

$$F1 = 2 * \frac{precision * recall}{precision + recall}$$

In der ROC-Kurve wird die Sensitivität auf der Ordinate und die Falsch-Positiv-Rate auf der Abszisse aufgetragen, die Fläche unter dieser Kurve kann dann auch im maschinellen Lernen als Gütekriterium herangezogen werden.

Bei Klassifikationsaufgaben, bei denen die Zielvariable nicht binär ist, müssen die o.g. Metriken angepasst werden. Dazu existieren verschiedene Ansätze:

- „Mikro“: Berechnet die Metriken global, indem die Summe der Richtig-Positiven, Falsch-Negativen und Falsch-Positiven gezählt wird.
- „Makro“: Berechnet die Metrik für jeden möglichen Wert der Zielvariablen und bildet dann deren ungewichteten Mittelwert. Dabei wird das Ungleichgewicht zwischen den einzelnen Klassen nicht berücksichtigt.
- „Gewichtet“: Berechnet die Metrik ebenfalls für jede mögliche Ausprägung der Zielvariablen, gewichtet diese dann aber anhand der Anzahl der wahren Instanzen für jedes Label. Durch diese Berechnungsvorschrift wird „Makro“ also dahingehend abgeändert, dass Unausgewogenheiten zwischen den jeweiligen Klassen/Merkmalausprägungen berücksichtigt werden. Zu beachten ist dabei, dass dieses Vorgehen zu einem Score führen kann, der nicht zwischen den Werten von Precision und Recall liegt.

3.2.2 Umsetzung/Implementierung

Zur Implementierung der maschinellen Lernverfahren wurden Python 3.7.5, Keras 2.3.1 und Scikit-Learn 0.22.3 [85] angewendet.

Python ist eine Programmiersprache, die häufig im Bereich der Datenverarbeitung und -analyse benutzt wird. Keras und Scikit-Learn sind open-source Software-Bibliotheken, in denen viele Algorithmen für maschinelles Lernen implementiert sind. Keras wird häufig für Deep Learning mit neuronalen Netzen verwendet, und Scikit-Learn enthält diverse Algorithmen für Klassifikations-, Regressions- und Clusteringaufgaben.

3.2.2.1 Vorbereitung der Daten

Wie bereits beschrieben, wurden fehlende kategoriale Werte bei allen Patienten durch den Wert -1 ersetzt, fehlende numerische Werte wurden aus dem Klinikinformationssystem extrahiert bzw. durch den Mittelwert der Werte der übrigen Patienten ersetzt.

Da viele maschinelle Lernverfahren nicht in der Lage sind, auf kategorialen Variablen zu arbeiten, müssen diese zunächst geeignet kodiert werden. Dazu wurde hier die sogenannte 1-aus-n- oder One-

Hot-Kodierung genutzt. Dabei wird jedes kategoriale Merkmal mit n möglichen Merkmalsausprägungen in einen binären Vektor der Länge n umgewandelt, der genau an einer Stelle den Wert 1 und an allen anderen den Wert 0 besitzt. Die Position der Zahl 1 repräsentiert die Ausprägung des Merkmals.

Die meisten maschinellen Lernverfahren sind zudem nicht skaleninvariant, sodass die numerischen Variablen geeignet skaliert werden müssen. Ziel ist es, dass die Daten einen Mittelwert von 0 und eine Standardabweichung von 1 haben. Dazu wurde die folgende Formel verwendet:

$$x_{scaled} = \frac{x - \bar{x}}{\sigma(x)} \quad (\bar{x}: \text{Mittelwert}, \sigma(x): \text{Standardabweichung})$$

3.2.2.2 Clusteranalyse

Da die in der vorliegenden Arbeit genutzte Permutation Feature Importance durch korrelierte Parameter stark verzerrt wird, wurde zunächst eine Clusteranalyse der numerischen Variablen basierend auf dem Spearman-Rangkorrelationskoeffizienten durchgeführt. Mit der Ward-Methode wurde dafür ein hierarchisches agglomeratives Clusterverfahren genutzt, bei dem die Cluster stets so fusioniert werden, dass der Zuwachs der totalen Varianz minimal ist [116]. Das zugehörige Dendrogramm ist in Abbildung 2 zu sehen. Um die Parameter zu Clustern zusammenzufügen, wurde anhand des Dendrogramms manuell ein Schwellenwert von 0,8 gewählt (gestrichelte Linie in Abbildung 2). Pro Cluster wurde ein Parameter selektiert, die restlichen Parameter wurden aus der weiteren Analyse ausgeschlossen. Die für die weitere Analyse ausgewählten Parameter sind in der Tabelle in Abbildung 2 fett markiert.

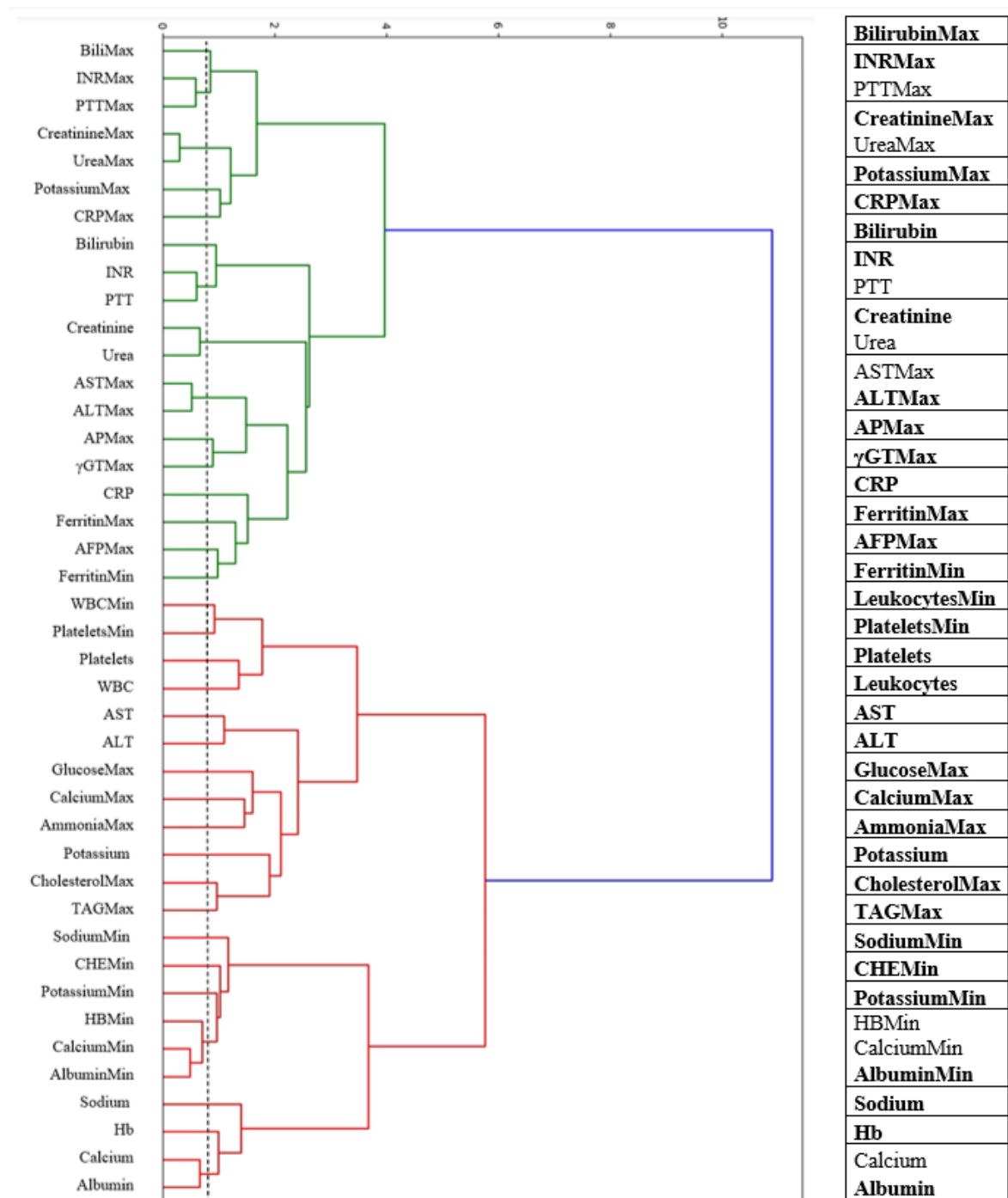


Abbildung 2: Dendrogramm Clusteranalyse

Ein Ausschnitt des zugehörigen Programmcodes ist im Folgenden zu sehen:

''' Die folgende Methode dient der Durchführung der Clusteranalyse. Dazu wird zunächst der Spearman-Rangkorrelationskoeffizient berechnet, darauf basierend wird dann eine hierarchische Clusteranalyse mit der Ward-Minimum-Varianz-Methode durchgeführt.

Die Ergebnisse werden in einem Dendrogramm visualisiert. '''

```
def clusterAnalysis(namesNumerical, numericalVar):  
    fig, axis1 = plt.subplots()  
    correl = spearmanr(numericalVar).correlation  
    correl_linkage = hierarchy.ward(correl)  
    dendrogr = hierarchy.dendrogram(correl_linkage, labels=numNames,  
    leaf_font_size=12, ax=axis1, leaf_rotation=90)  
    dendrogr_idx = np.arange(0, len(dendrogr['ivl']))  
    plt.show()
```

3.2.2.3 Ablauf des Trainierens und Testens

Nach der Vorbereitung der Daten (Kodierung und Skalierung) und Durchführung der Clusteranalyse wurde das resultierende Datenset zufällig in 85 % Trainingsdaten und 15 % Testdaten gesplittet und auf den Trainingsdaten wurde dann jeweils ein Modell trainiert. Dabei wurden die Hyperparameter des Modells mit Hilfe einer crossvalidierten Grid Search [9] optimiert, als Kriterium wurde hierfür die Accuracy verwendet.

Als zu optimierende Hyperparameter wurden beim Entscheidungsbaum das Auswahlkriterium (Gini-Koeffizient vs. Entropie), die minimale Anzahl an Trainingsbeispielen pro Knoten (5, 10, 15, 20, 25, 50) sowie die maximale Tiefe des Baums (3, 4, 5, 6) gewählt, bei den Random Forests zusätzlich die Anzahl an Bäumen pro Ensemble (10, 100, 500). Bei der SVM wurden die Auswahl der Kernelfunktion (radiale Basisfunktion, polynomial, sigmoidal), sowie die Kernelparameter γ (1e-1, 1e-2, 1e-3, 1e-4, 1e-5) und C (0,00001; 0,000001; 0,001; 0,0001; 0,01; 0,1; 1; 10; 100) optimiert, bei den neuronalen Netzen die Anzahl an Neuronen in den versteckten Schichten sowie die Aktivierungsfunktion.

Im nächsten Schritt wurde das trainierte Modell auf die Testdaten angewendet. Zur Beurteilung der Performance wurden die Accuracy, der F1-Score sowie die AUC der ROC-Kurve bestimmt. Außerdem wurde die PFI analysiert. Ob die PFI auf die Trainings- oder Testdaten angewendet werden soll, ist noch Gegenstand der Forschung, hier wurde sie für eine möglichst gute Generalisierbarkeit auf die Testdaten angewendet.

Damit die PFIs möglichst unabhängig von einem konkreten Modell sind, wurde der Prozess des zufälligen Splittens der Datenbank, des Trainierens und Testens inkl. Hyperparameteroptimierung sowie der Anwendung der PFI für jedes Modell 50 Mal wiederholt und der Mittelwert der Ergebnisse gebildet.

Exemplarisch ist im Folgenden ein Teil des Programmcodes für Entscheidungsbäume abgebildet:

```
''' Als Klassifizierer wird hier exemplarisch ein Entscheidungsbaum  
gewählt. Die verschiedenen Hyperparameter (z. B. Wahl des  
Entscheidungskriteriums, minimale Anzahl an Beispielen pro Blatt und  
maximale Tiefe des Baums) werden in einer cross-validierten Gridsearch  
getestet, die beste Hyperparameterkombination (gemessen an der Accuracy  
des entstehenden Modells) wird ausgewählt und anschließend zum Training  
des Modells verwendet. Der entstehende Klassifizier wird auf die Testdaten  
angewendet'''
```

```
tuned_parameters = [{'criterion': ['gini', 'entropy'], 'max_depth': [3, 4,  
5, 6], 'min_samples_leaf': [5, 10, 15, 20, 25, 50]}]
```

```
clf = GridSearchCV(DecisionTreeClassifier(), tuned_parameters,  
scoring='accuracy')
```

```
# Training und Anwendung des Classifiers auf den Trainingsdaten  
clf.fit(X_train, y_train)  
y_predTR = clf.predict(X_train)
```

```
# Anwenden der Classifiers auf die Testdaten  
y_true, y_pred = y_test, clf.predict(X_test)
```

```
# Anwenden der Permutation Feature Importance auf die Testdaten  
results = permutation_importance(clf, X_test, y_test, scoring='accuracy')
```

```
# Bestimmung der Accuracy, der ROC-AUC und des F1-Scores  
return results.importances_mean, results.importances_std,  
accuracy_score(y_train, y_predTR), accuracy_score(y_test, y_pred),  
roc_auc_score(y_train, y_scoreTR), roc_auc_score(y_test, y_score),  
f1_score(y_train, y_predTR, average="macro"), f1_score(y_test, y_pred,  
average="macro")
```

4 Ergebnisse

Die INCA-Trial-Datenbank umfasst Daten von insgesamt 983 Patienten, davon waren 65,8 % Männer. Die häufigste Ätiologie im Studienkollektiv war chronischer Alkoholkonsum (54 %), gefolgt von einer chronischen Hepatitis C Virus-Infektion (13,6 %). Zum Indexzeitpunkt hatten 313 Patienten (31,8 %) noch nie eine Dekompensation der Zirrhose erlitten, 354 Patienten (36 %) hatten bereits ein Dekompensationsereignis in der Vergangenheit, waren aber zum gegenwärtigen Zeitpunkt rekompensiert, und 316 Patienten (32,1 %) waren im dekompenzierten Stadium. 825 Patienten (83,9 %) nahmen am Follow-up teil. Die mittlere Follow-up Dauer betrug 12 Monate, die maximale 55 Monate. 46,5 % der Patienten, die am Follow-up teilnahmen, dekompenzierten während dieser Zeit.

Wie in Tabelle 4 und Tabelle 5 zu sehen ist, erreichten Random Forests bei der retrospektiven Einschätzung einer Dekompensation die höchste Accuracy, bei der prospektiven Vorhersage zeigten SVM die besten Ergebnisse.

Tabelle 4: Accuracy der einzelnen Verfahren zur retrospektiven Beurteilung einer Dekompensation

Frühere Dekompensation	Accuracy Trainingsdaten	Accuracy Testdaten
Decision Trees	0,719	0,655
Random Forests	0,816	0,705
SVM	0,784	0,685
Neuronale Netze	0,916	0,65

Tabelle 5: Accuracy bezüglich der prospektiven Einschätzung einer Dekompensation

Dekompensation während Follow-up	Accuracy Trainingsdaten	Accuracy Testdaten
Decision Trees	0,786	0,711
Random Forests	0,826	0,693
SVM	0,786	0,738
Neuronale Netze	0,823	0,730

Auf die Ergebnisse der einzelnen Verfahren wird im Folgenden genauer eingegangen, die vollständigen Tabellen mit allen PFI-Werten befinden sich im Anhang.

4.1 Entscheidungsbäume

4.1.1 Retrospektive Beurteilung

Ungeclustert:

Wie in Tabelle 6 zu sehen ist, erreichten Entscheidungsbäume auf den ungeclusterten Trainingsdaten zur Vorhersage einer früheren Dekompensation eine Accuracy von 0,72, eine AUC von 0,87 und einen F1-Score von 0,72. Bezogen auf die Testdaten betrugen die Accuracy 0,66, die AUC 0,81 und der F1-Score 0,66.

Tabelle 6: Metriken Entscheidungsbaum ungeclustert retrospektiv

	Trainingsdaten	Testdaten
Accuracy	0,718	0,657
AUC	0,872	0,808
F1	0,716	0,655

Abbildung 3 stellt die Top 20 Parameter basierend auf der PFI dar. Der Baselinwert des Albumins wurde mit Abstand am höchsten gewichtet, gefolgt vom maximalen Bilirubinwert und fehlenden Informationen über *NOD2*-Genotypen.

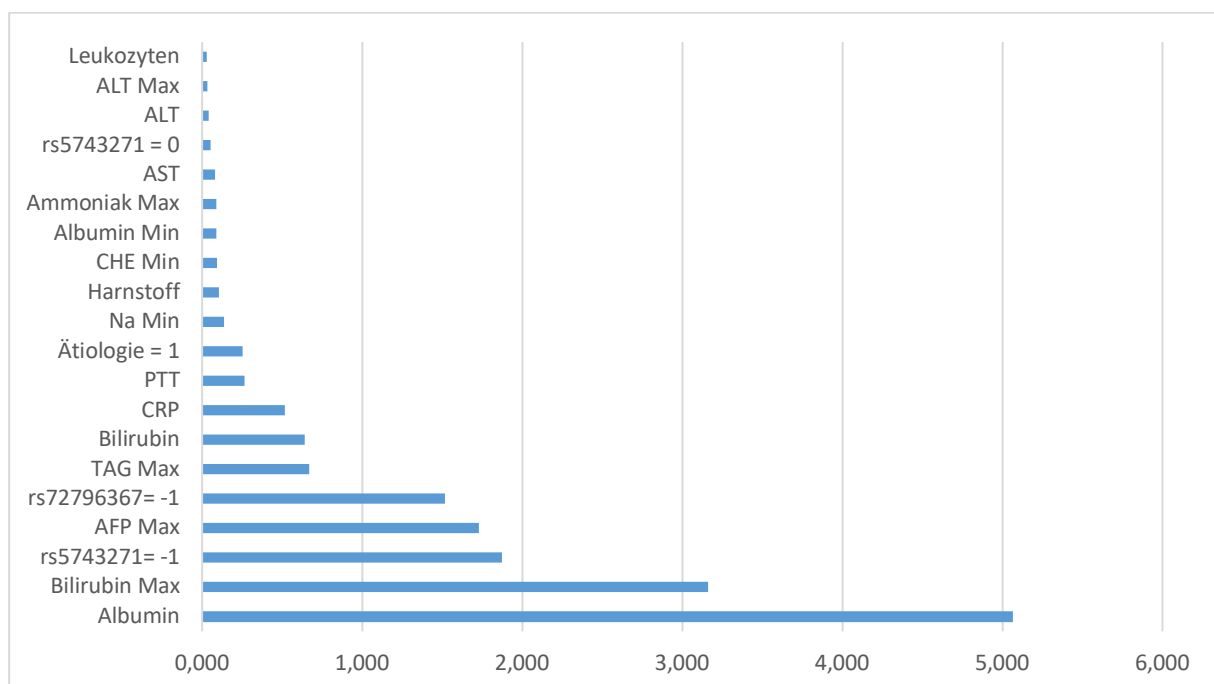


Abbildung 3: Top 20 Entscheidungsbaum ungeclustert retrospektiv

Geclustert:

Tabelle 7 zeigt, dass Entscheidungsbäume auf den geclusterten Trainingsdaten eine Accuracy von 0,72, eine AUC von 0,87 und einen F1-Score von 0,72 erreichten, auf den Testdaten betrugen die entsprechenden Werte 0,66, 0,8 und 0,65.

Tabelle 7: Metriken Entscheidungsbaum geclustert retrospektiv

	Trainingsdaten	Testdaten
Accuracy	0,719	0,655
AUC	0,872	0,802
F1	0,718	0,653

In Abbildung 4 sind die 20 wichtigsten Parameter zur retrospektiven Vorhersage einer früheren Dekompensation durch Entscheidungsbäume zu sehen. Wie auch im ungeclusterten Datenset werden der Baselinewert des Albumins und der maximale Bilirubinwert am höchsten gewichtet, nun aber gefolgt vom maximalen Wert des AFP.

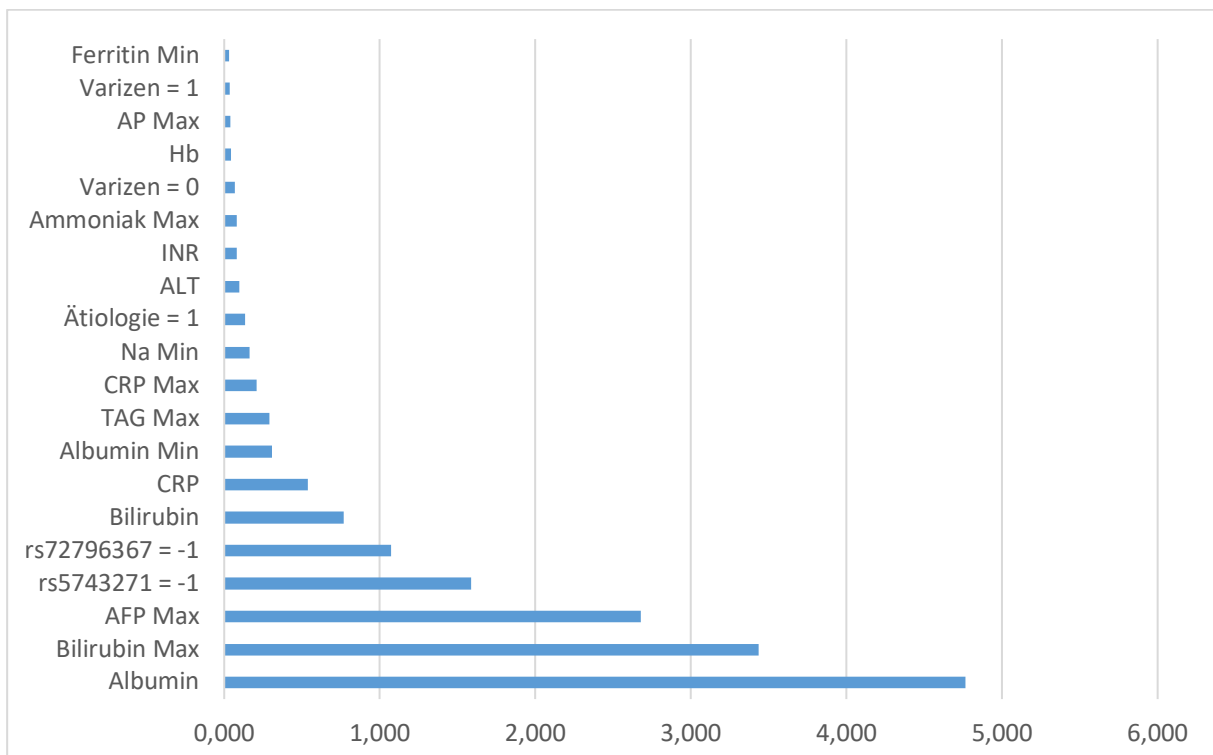


Abbildung 4: Top 20 Entscheidungsbaum geclustert retrospektiv

4.1.2 Prospektive Beurteilung

Ungeclustert:

Bezogen auf die prospektive Vorhersage einer Dekompensation in der Follow-up Periode erreichten Entscheidungsbäume auf den Trainingsdaten eine Accuracy von 0,78, eine AUC von 0,85 und einen F1-Score von 0,78, bezogen auf die Testdaten betrugen die Werte 0,71, 0,78 und 0,7 (Tabelle 8).

Tabelle 8: Metriken Entscheidungsbaum ungeclustert prospektiv

	Trainingsdaten	Testdaten
Accuracy	0,779	0,707
AUC-ROC	0,852	0,775
F1	0,777	0,702

Abbildung 5 zeigt die Top 20 Parameter, die aus dem ungeclusterten Datenset zur prospektiven Vorhersage einer Dekompensation ermittelt wurden. Durch die PFI wurden die maximalen Werte von Bilirubin und Harnstoff sowie der Baselinewert von Albumin und der minimale Wert der CHE am höchsten gewichtet.

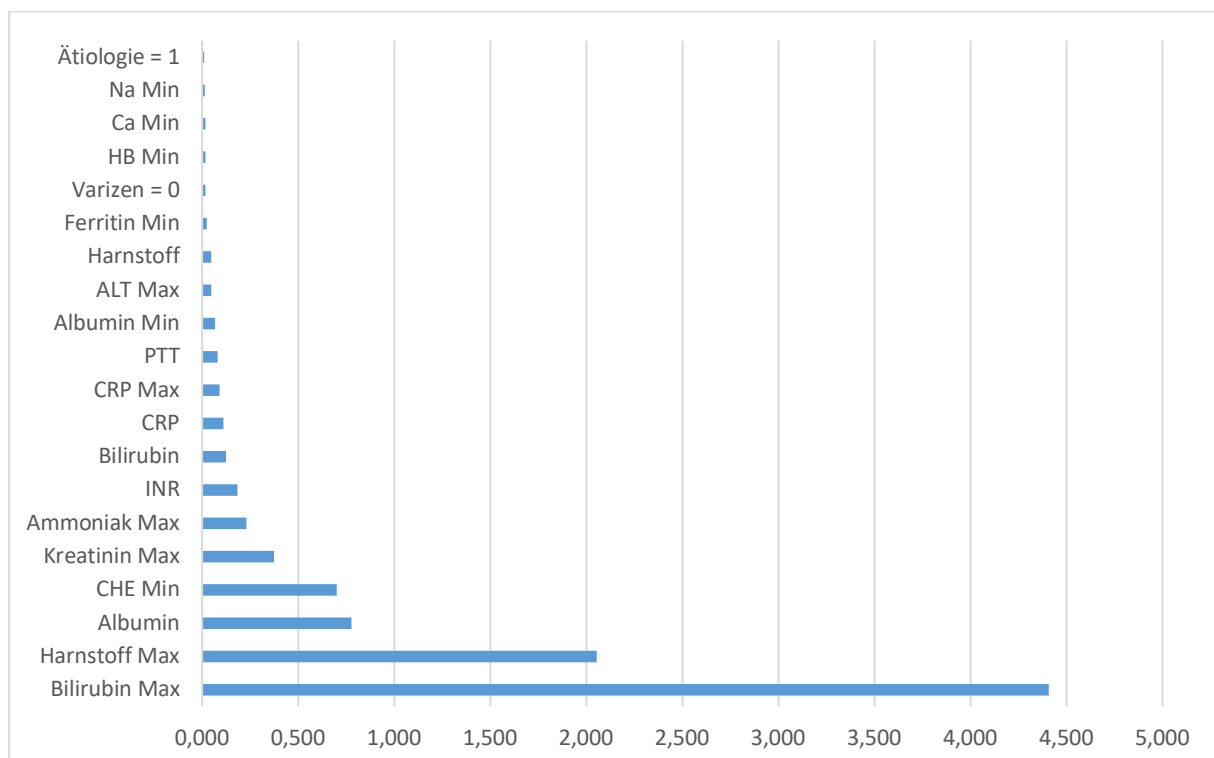


Abbildung 5: Top 20 Entscheidungsbaum ungeclustert prospektiv

Gelclustert:

Tabelle 9 listet die Metriken zur prospektiven Vorhersage der Dekompensation basierend auf den gelclusterten Daten auf. Auf den Trainingsdaten betrugen die Accuracy 0,79, die AUC 0,86 und der F1-Score 0,78, auf den Testdaten 0,71, 0,77 und 0,71.

Tabelle 9: Metriken Entscheidungsbaum gelclustert prospektiv

	Trainingsdaten	Testdaten
Accuracy	0,786	0,711
AUC-ROC	0,858	0,770
F1	0,784	0,709

Wie in Abbildung 6 zu sehen ist, wurden die maximalen Werte von Bilirubin und Kreatinin sowie der Baselinewert von INR und der minimale Wert der CHE auf dem gelclusterten Datenset am höchsten gewichtet.

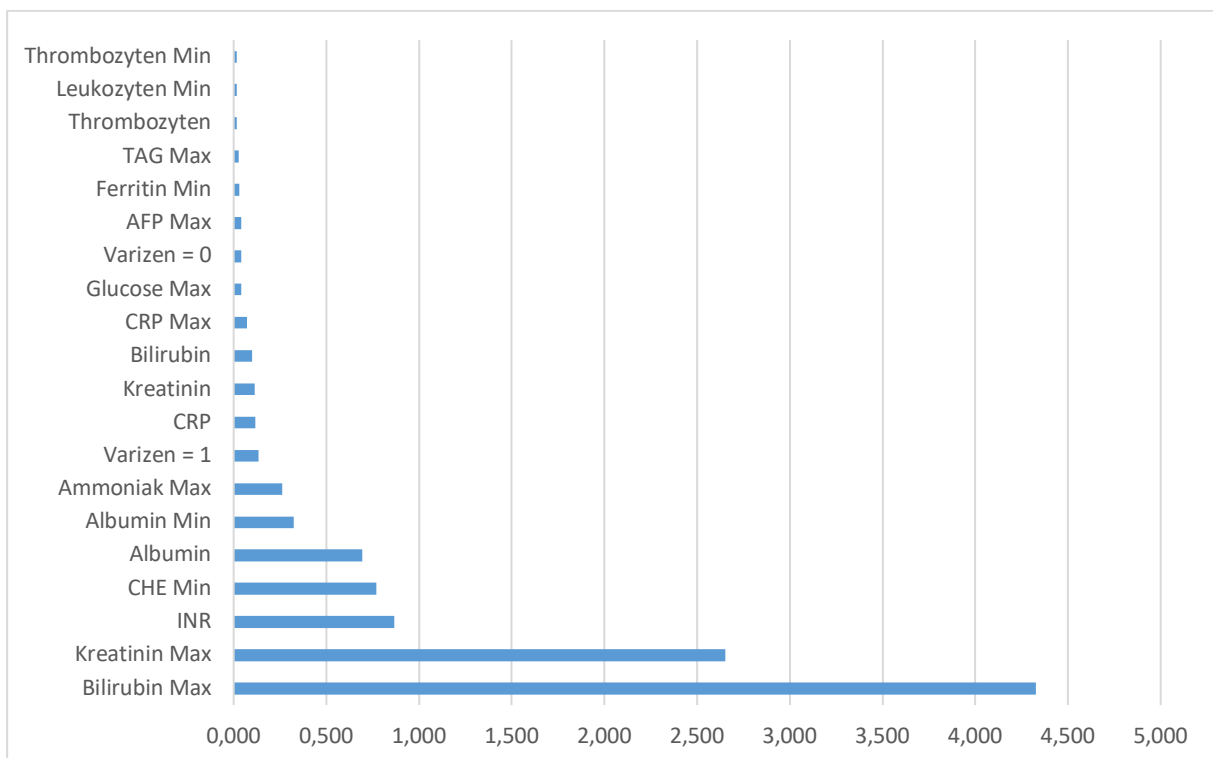


Abbildung 6: Top 20 Entscheidungsbaum gelclustert prospektiv

4.2 Random Forests

4.2.1 Retrospektive Beurteilung

Ungeclustert:

In Tabelle 10 ist zu sehen, dass Random Forests auf den ungeclusterten Trainingsdaten zur Vorhersage einer früheren Dekompensation eine Accuracy von 0,82, eine AUC von 0,96 und einen F1-Score von 0,82 erreichten. Bezogen auf die Testdaten betrugen die entsprechenden Werte 0,69, 0,86 und 0,68.

Tabelle 10: Metriken Random Forest ungeclustert retrospektiv

	Trainingsdaten	Testdaten
Accuracy	0,820	0,687
AUC	0,956	0,860
F1	0,819	0,682

In Abbildung 7 sind die Top 20 Parameter des Random Forests basierend auf der PFI dargestellt. Am höchsten wurde der Maximalwert des Bilirubins gewichtet, gefolgt von den Baselinewerten von Albumin, Bilirubin und CRP.

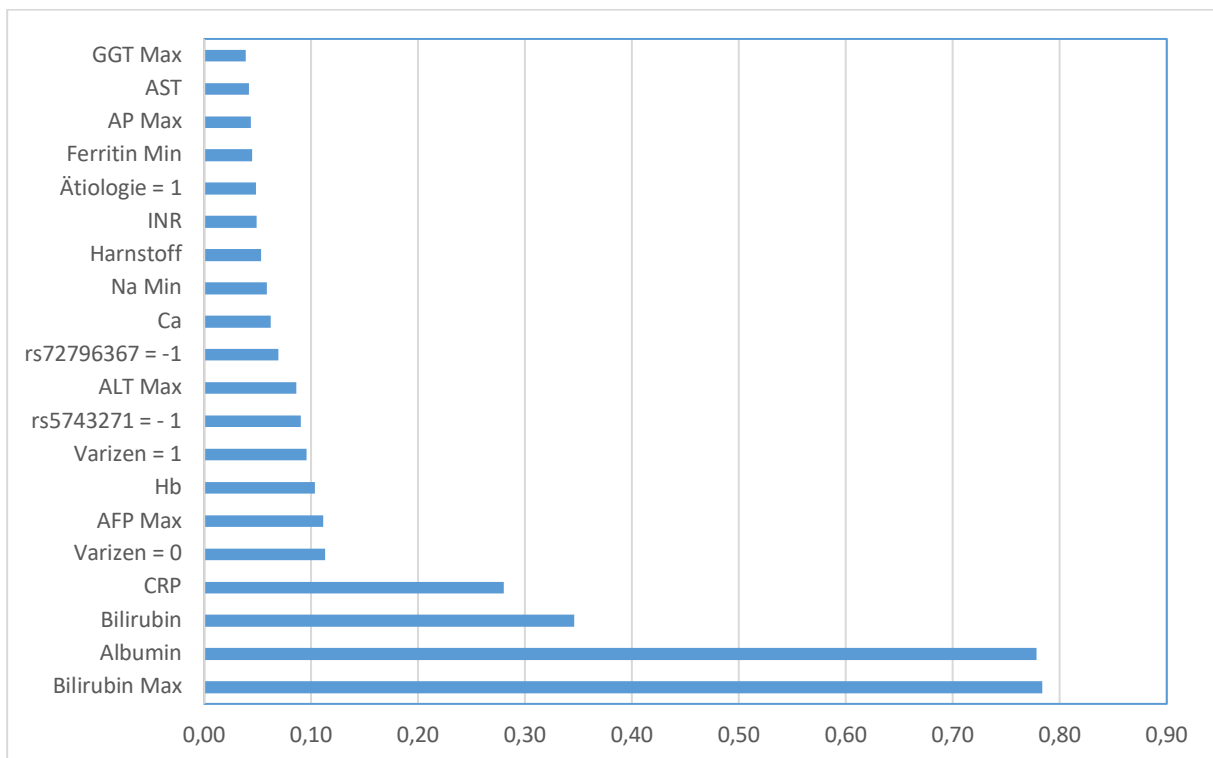


Abbildung 7: Top 20 Random Forest ungeclustert retrospektiv

Geclustert:

Auf dem geclusterten Datenset erreichten Random Forests näherungsweise die gleichen Werte wie auf den ungeclusterten Parametern (Tabelle 11).

Tabelle 11: Metriken Random Forest geclustert retrospektiv

	Trainingsdaten	Testdaten
Accuracy	0,816	0,705
AUC	0,955	0,869
F1	0,815	0,701

Die 20 wichtigsten Parameter auf dem geclusterten retrospektiven Datenset sind in Abbildung 8 dargestellt. Durch die PFI wurde der Baselinewert vom Albumin am höchsten gewichtet, dicht gefolgt vom maximalen Bilirubinwert. Die Baselinewerte von Bilirubin und CRP wurden ebenfalls als relevant eingeschätzt. Informationen über eine langfristige Antibiotikaeinnahme befanden sich auch unter den Top 20 Parametern.

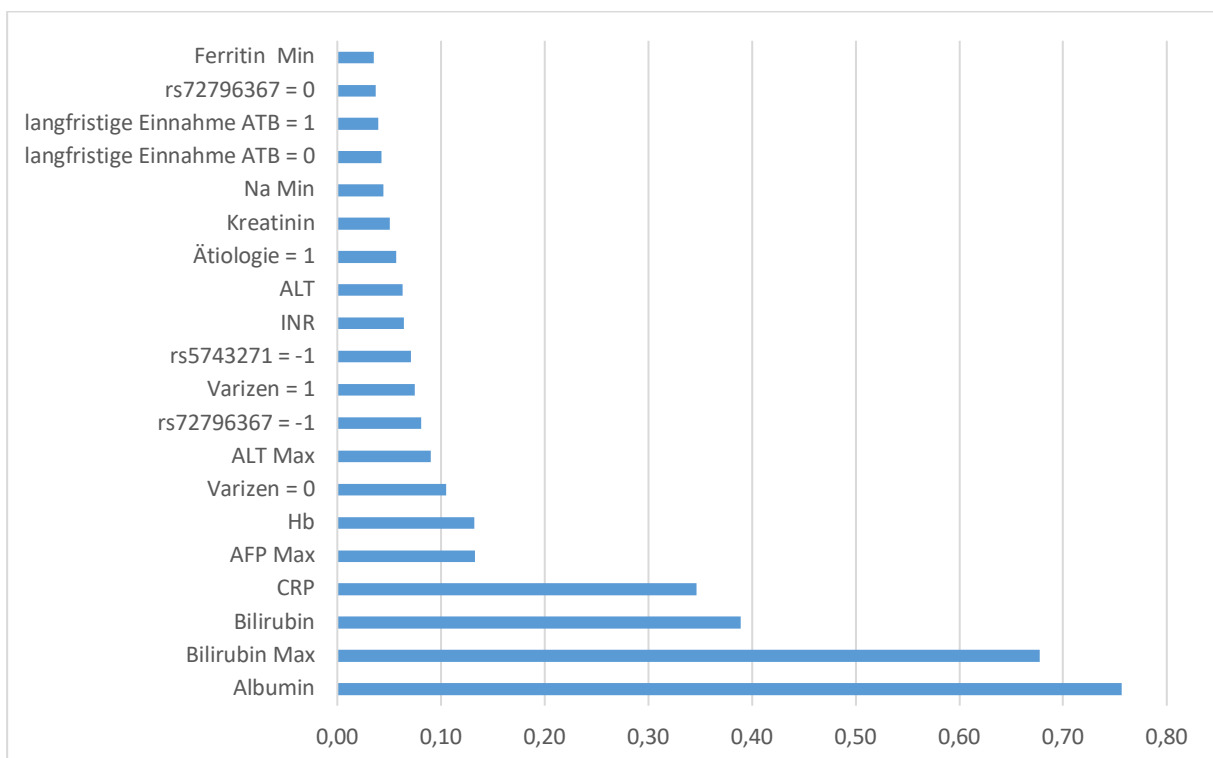


Abbildung 8: Top 20 Random Forest geclustert retrospektiv

4.2.2 Prospektive Beurteilung

Ungeclustert:

Bezüglich der prospektiven Vorhersage einer Dekompensation in der Follow-up-Periode erreichten Random Forests auf dem ungeclusterten Datenset eine Accuracy von 0,82, eine AUC von 0,9 und einen F1-Score von 0,82. Die entsprechenden Werte auf den Testdaten betrugen 0,75, 0,83 und 0,74 (Tabelle 12).

Tabelle 12: Metriken Random Forest ungeclustert prospektiv

	Trainingsdaten	Testdaten
Accuracy	0,817	0,745
AUC	0,900	0,825
F1	0,816	0,743

Wie in Abbildung 9 gezeigt wird, wurden die maximalen Bilirubin- und Kreatininwerte sowie die Baselinewerte von Albumin, Bilirubin und CRP hoch gewichtet.

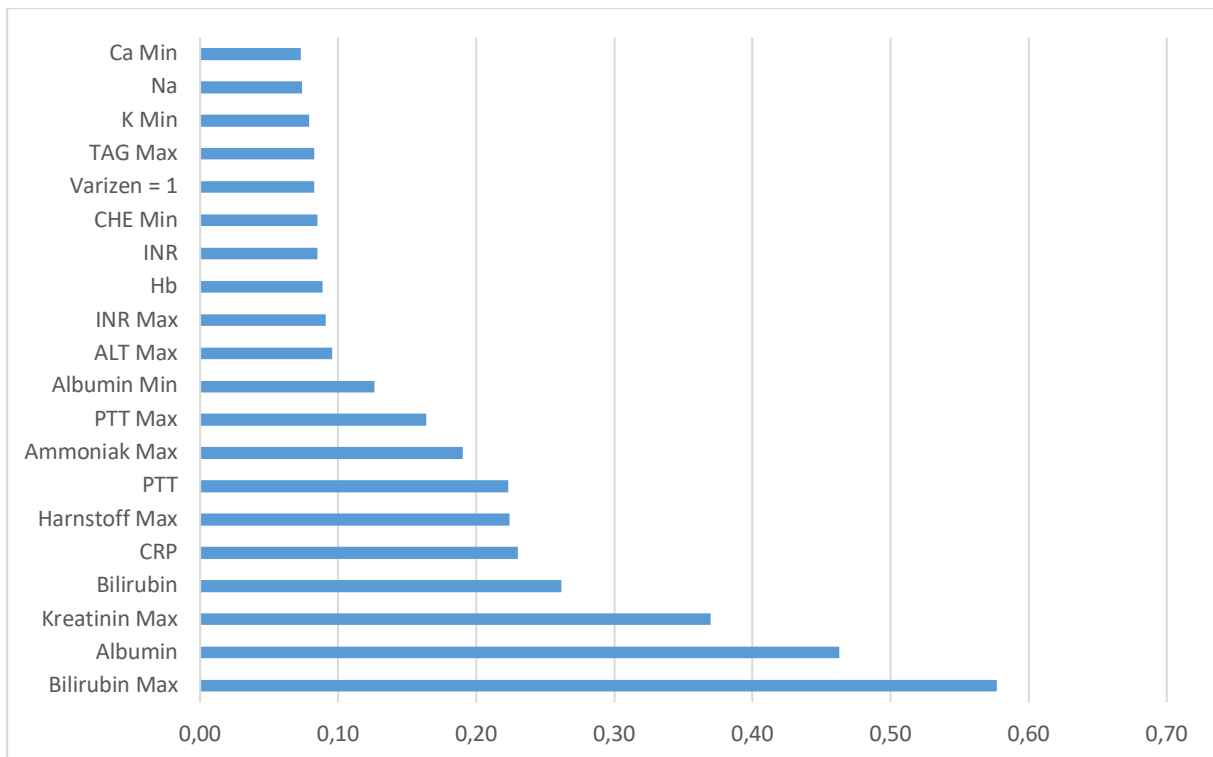


Abbildung 9: Top 20 Random Forest ungeclustert prospektiv

Geklustert:

Tabelle 13 zeigt die Metriken für das geklusterte Datenset. Auf den Trainingsdaten betrugen die Accuracy 0,83, die AUC 0,96 und der F1-Score 0,83, auf den Testdaten 0,69, 0,86 und 0,69.

Tabelle 13: Metriken Random Forest geklustert prospektiv

	Trainingsdaten	Testdaten
Accuracy	0,826	0,693
AUC	0,962	0,861
F1	0,826	0,688

Wie auch im ungeklusterten Datenset wurde der maximale Wert des Bilirubins am höchsten gerankt. Es folgten der Maximalwert des Kreatinins und die Baselinewerte von Albumin, CRP und Bilirubin (Abbildung 10).

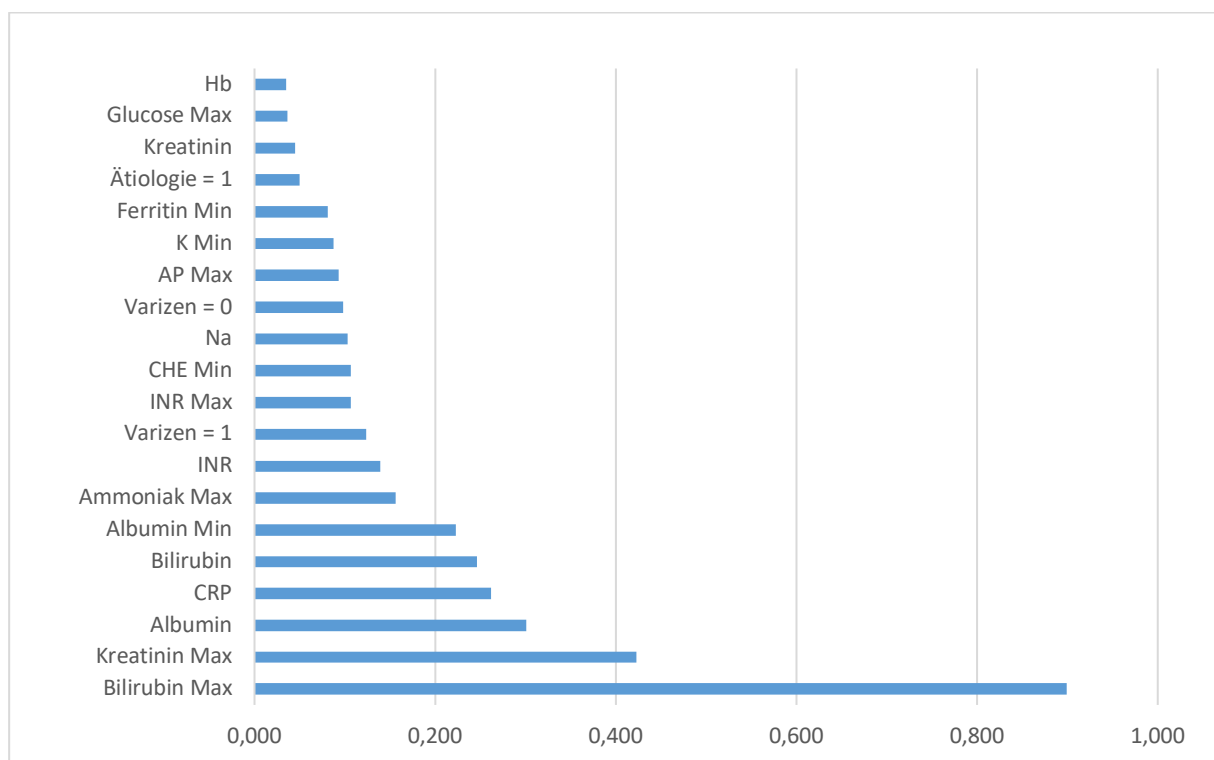


Abbildung 10: Top 20 Random Forest geklustert prospektiv

4.3 Support Vector Machine

4.3.1 Retrospektive Beurteilung

Ungeclustert:

Für die retrospektive Einschätzung einer Dekompensation erreichten SVM auf dem ungeclusterten Datenset eine Accuracy von 0,8, eine AUC von 0,92 und einen F1-Score in Höhe von 0,8 auf den Trainingsdaten. Auf den Testdaten betrugen die entsprechenden Werte 0,68, 0,85 und 0,68 (Tabelle 14).

Tabelle 14: Metriken SVM ungeclustert retrospektiv

	Trainingsdaten	Testdaten
Accuracy	0,795	0,682
AUC	0,918	0,846
F1	0,795	0,681

Für die SVM waren auf dem ungeclusterten Datenset die Baselinewerte von Bilirubin und Albumin gefolgt von den Minimalwerten von Natrium und Albumin am wichtigsten (Abbildung 11).

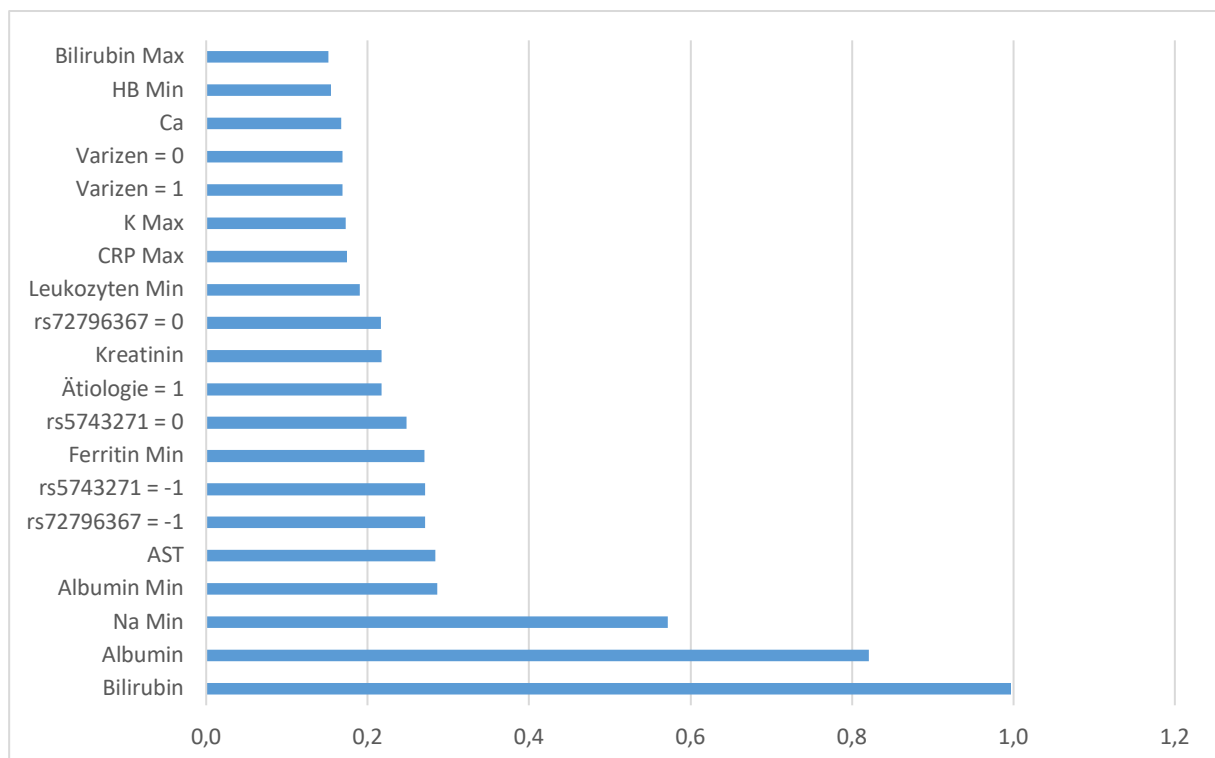


Abbildung 11: Top 20 SVM ungeclustert retrospektiv

Geclustert:

Bezogen auf die retrospektive Vorhersage einer Dekompensation auf geclusterten Trainingsdaten betrugen die Accuracy 0,78, die AUC 0,91 und der F1-Score 0,79. Auf den Testdaten erreichte die SVM eine Accuracy von 0,69, eine AUC von 0,85 und einen F1-Score von 0,68 (Tabelle 15).

Tabelle 15: Metriken SVM geclustert retrospektiv

	Trainingsdaten	Testdaten
Accuracy	0,784	0,685
AUC	0,913	0,851
F1	0,785	0,684

Wie auch auf den ungeclusterten Daten wurden die Baselinewerte von Bilirubin und Albumin sowie der minimale Natriumwert durch die PFI am höchsten gerankt. Es folgten der Baselinewert der AST und die *NOD2*-Genotypen (Abbildung 12).

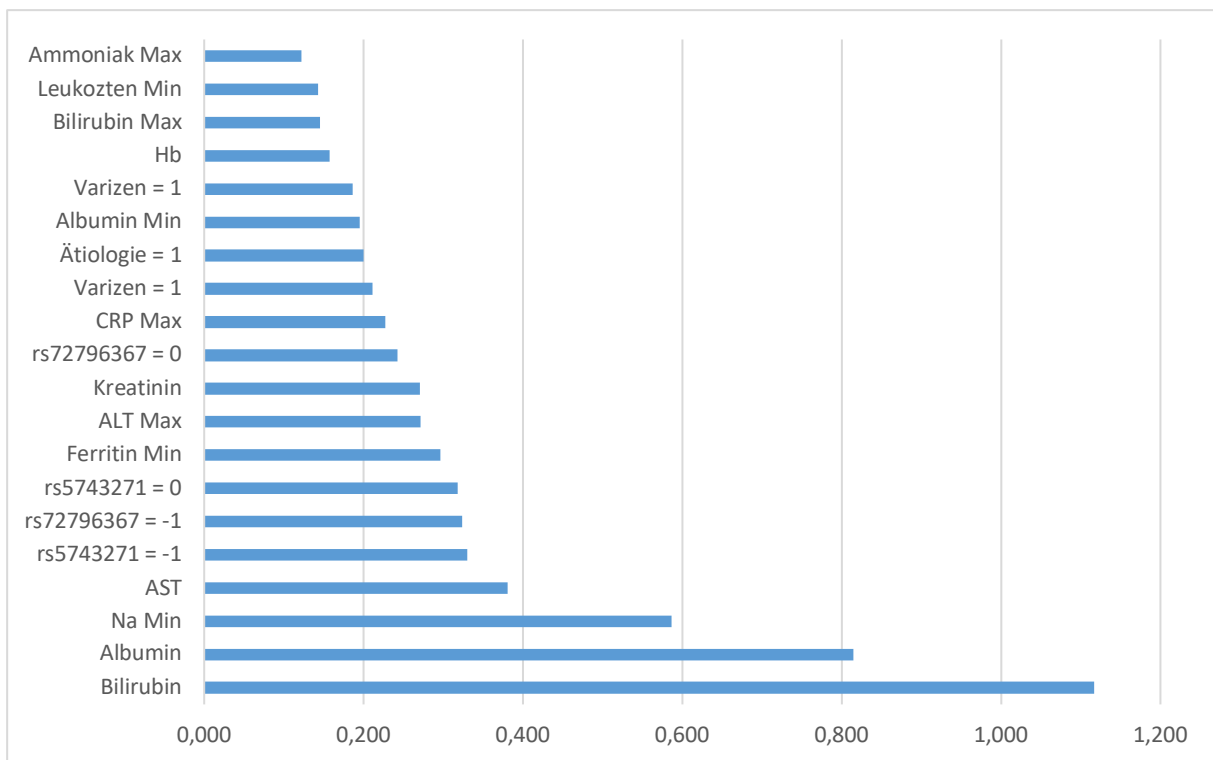


Abbildung 12: Top 20 SVM geclustert retrospektiv

4.3.2 Prospektive Beurteilung

Ungeclustert:

Bezüglich der prospektiven Vorhersage einer Dekompensation in der Follow-up-Periode erreichten SVM auf dem ungeclusterten Datenset eine Accuracy von 0,79, eine AUC von 0,87 und einen F1-Score von 0,79. Die entsprechenden Werte auf den Testdaten betrugen 0,73, 0,8 und 0,73 (Tabelle 16).

Tabelle 16: Metriken SVM ungeclustert prospektiv

	Trainingsdaten	Testdaten
Accuracy	0,791	0,732
AUC	0,867	0,801
F1	0,791	0,730

In Abbildung 13 ist zu sehen, dass der maximale Bilirubinwert mit Abstand am relevantesten war. Die Baselinewerte von Natrium, Albumin und PTT sowie der maximale Ammoniakwert wurden durch die PFI ebenfalls hoch gerankt.

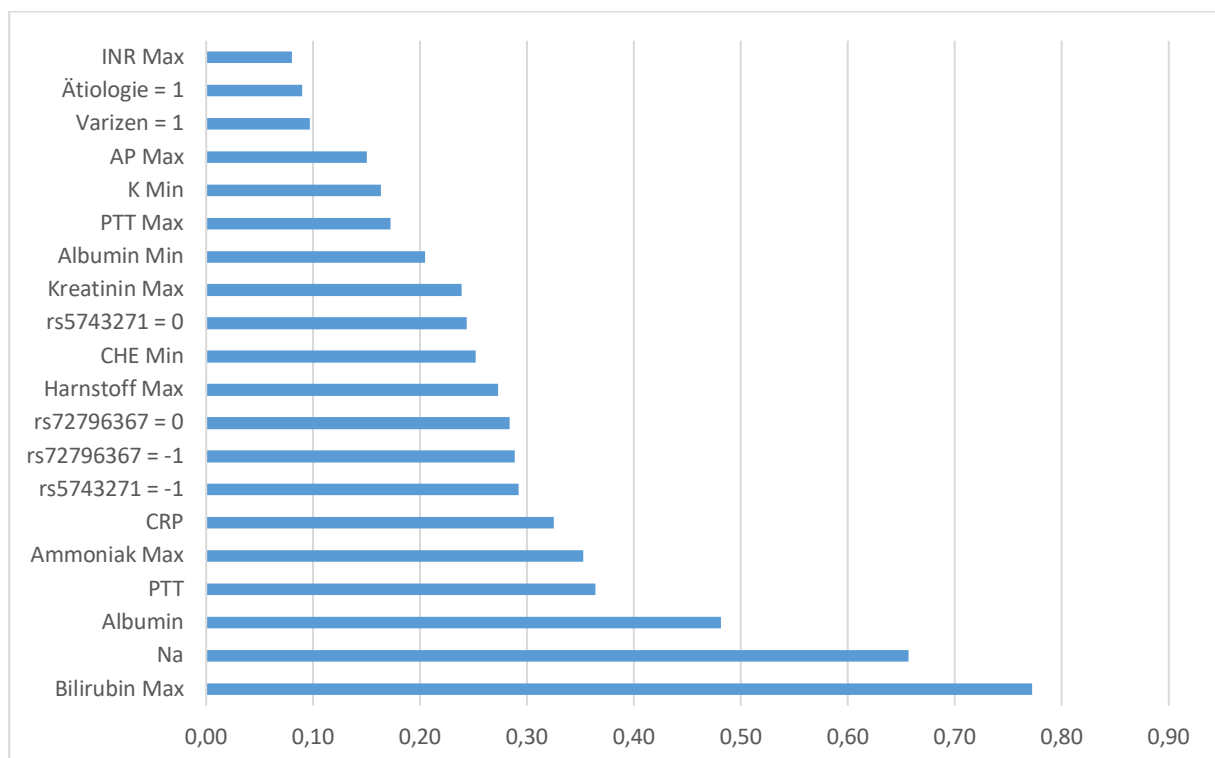


Abbildung 13: Top 20 SVM ungeclustert prospektiv

Gelclustert:

Tabelle 17 listet die Metriken zur prospektiven Vorhersage der Dekompensation durch SVM basierend auf den gelclusterten Daten auf. Auf den Trainingsdaten betrugen die Accuracy 0,79, die AUC 0,86 und der F1-Score 0,79, auf den Testdaten 0,74, 0,82 und 0,74.

Tabelle 17: Metriken SVM gelclustert prospektiv

	Trainingsdaten	Testdaten
Accuracy	0,786	0,738
AUC	0,857	0,817
F1	0,785	0,737

Bezogen auf die gelclusterten Daten war der maximale Bilirubinwert am wichtigsten, dicht gefolgt von den Baselinewerten von Natrium, Albumin und CRP (Abbildung 14). Die Einnahme von PPI und ein weibliches Geschlecht befanden sich ebenfalls unter den Top 20 Parametern.

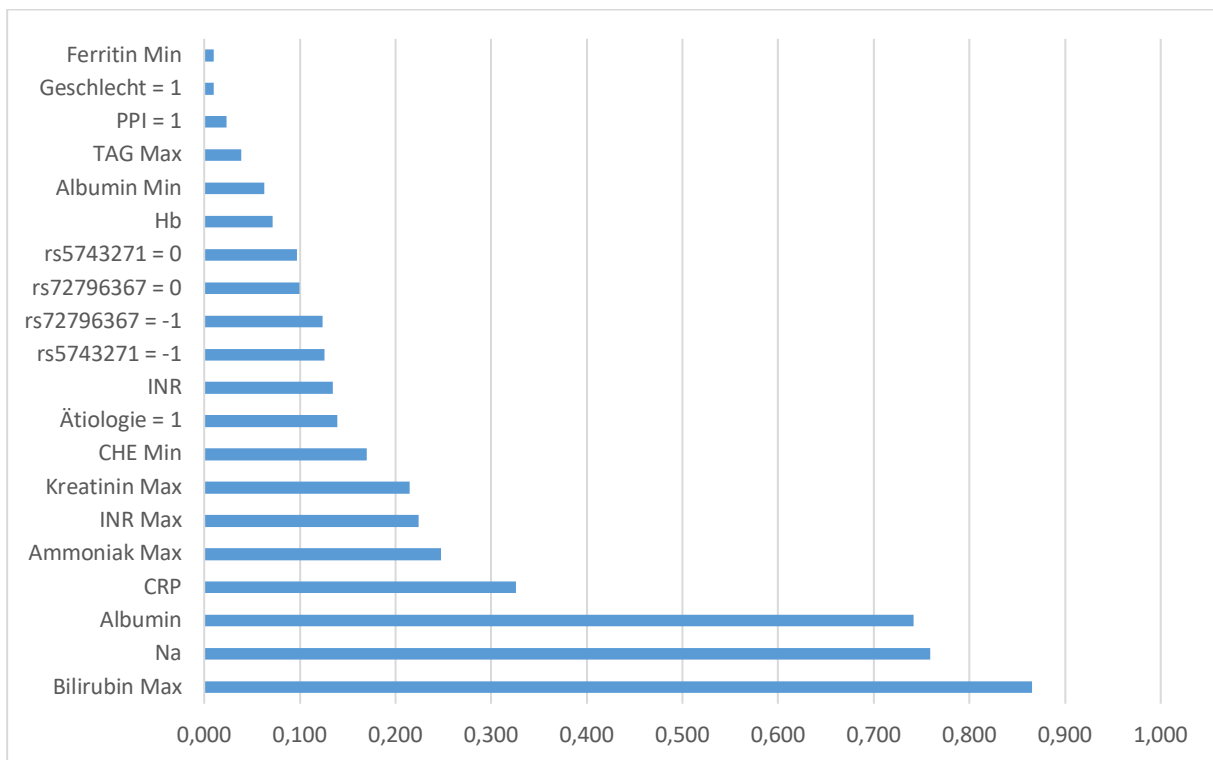


Abbildung 14: Top 20 SVM gelclustert prospektiv

4.4 Neuronale Netze

4.4.1 Retrospektive Beurteilung

Ungeclustert:

Neuronale Netze erreichten bei der retrospektiven Beurteilung einer Dekompensation auf dem ungeclusterten Trainingsdatensatz eine Accuracy von 0,91, eine AUC von 0,96 und einen F1-Score von 0,91, auf den Testdaten betrugen die entsprechenden Werte 0,65, 0,81 und 0,65 (Tabelle 18).

Tabelle 18: Metriken neuronales Netz ungeclustert retrospektiv

	Trainingsdaten	Testdaten
Accuracy	0,912	0,650
AUC	0,959	0,814
F1	0,910	0,648

Durch die PFI wurde der Baselinewert des Bilirubins mit Abstand am höchsten gewichtet, der Baselinewert des Albumins sowie die Minimalwerte von Natrium und Albumin waren ebenfalls relevant. Eine langfristige Antibiotikaeinnahme, die Calciumkonzentration im Serum und die minimale Leukozytenzahl befanden sich auch unter den Top 20 Parametern (Abbildung 15).

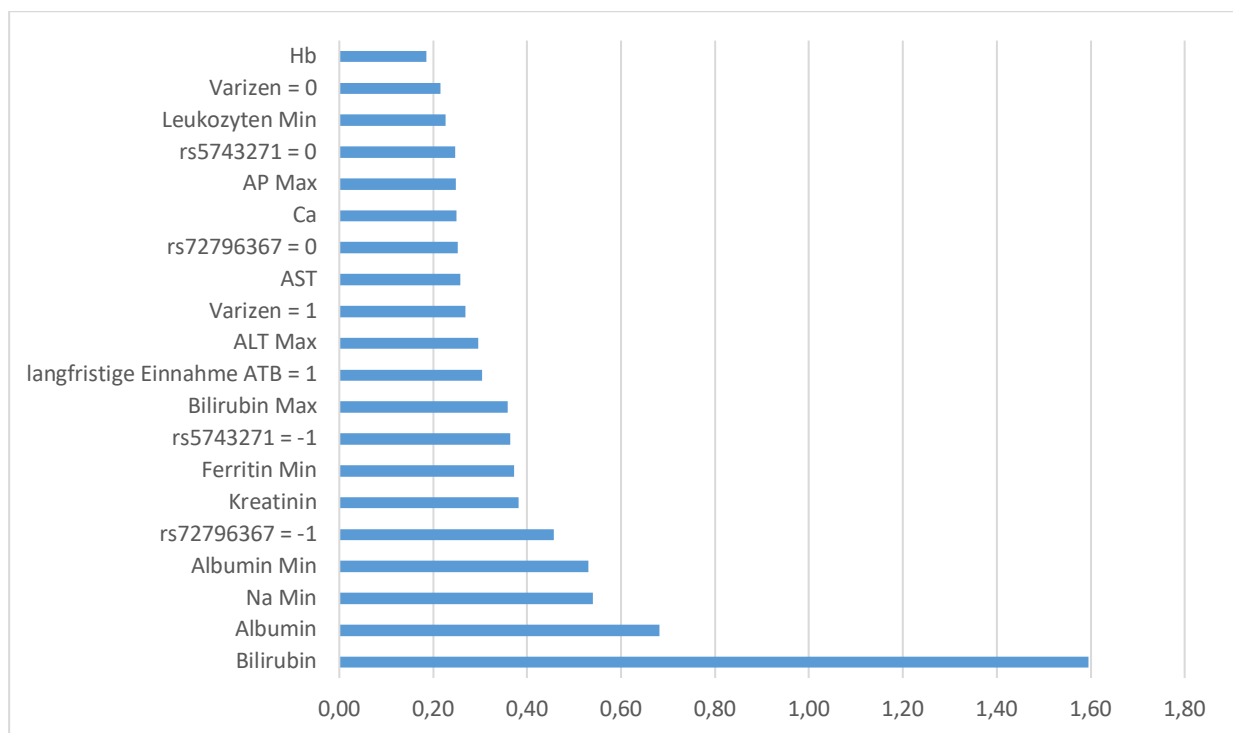


Abbildung 15: Top 20 neuronales Netz ungeclustert retrospektiv

Geclustert:

Auf dem geclusterten Trainingsdatensatz betrugen die Accuracy 0,92, die AUC 0,96 und der F1-Score 0,92 für die retrospektive Einschätzung einer Dekompensation, bezogen auf die Testdaten betrugen die entsprechenden Werte 0,65, 0,81 und 0,65 (Tabelle 19).

Tabelle 19: Metriken neuronales Netz geclustert retrospektiv

	Trainingsdaten	Testdaten
Accuracy	0,916	0,650
AUC	0,959	0,809
F1	0,916	0,650

Wie auch im ungeclusterten Datenset wurden die Baselinewerte von Bilirubin und Albumin sowie der minimale Natriumwert am höchsten gewichtet, hier jedoch gefolgt von dem Fehlen von Informationen über *NOD2*-Genotypen (Abbildung 16). Der minimale Ferritinwert wurde durch die PFI hoch gewichtet, zudem befanden sich Informationen über eine Antibiotikaeinnahme und die Kaliumwerte unter den Top 20 Parametern.

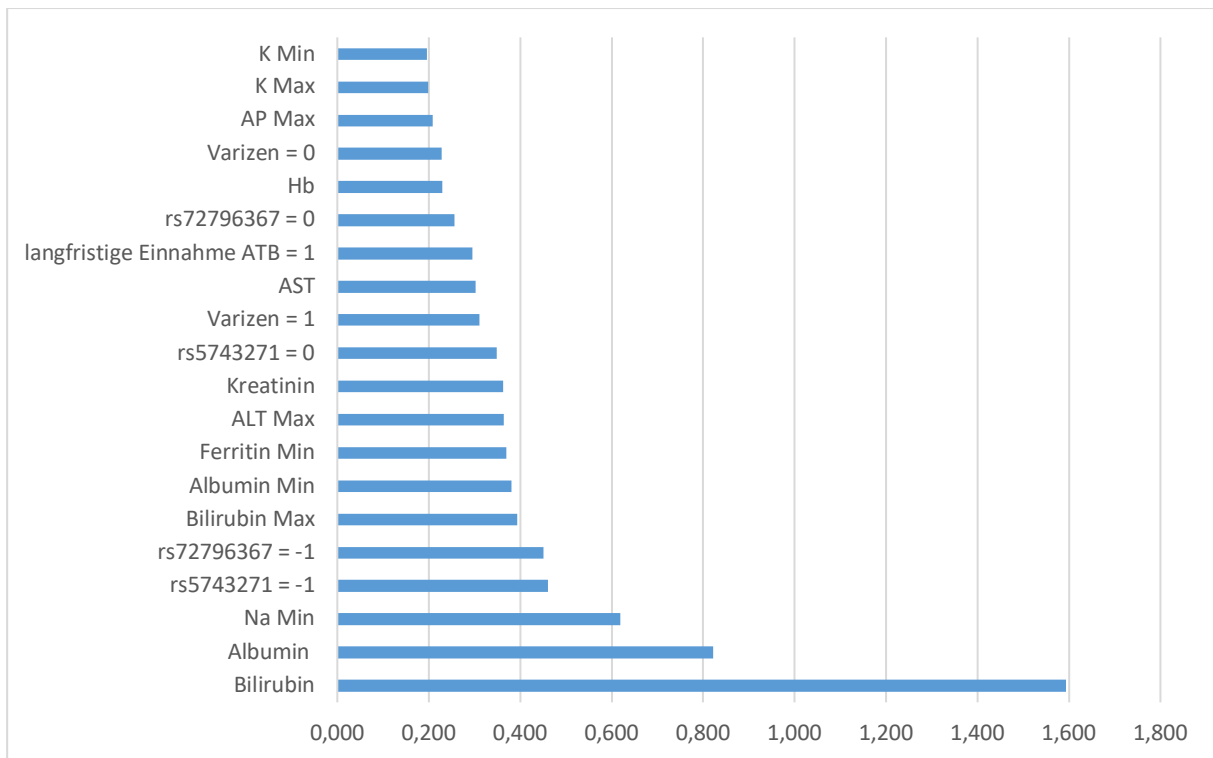


Abbildung 16: Top 20 neuronales Netz retrospektiv geclustert

4.4.2 Prospektive Beurteilung

Ungeclustert:

In Tabelle 20 ist zu sehen, dass neuronale Netze auf den ungeclusterten Trainingsdaten zur Vorhersage einer Dekompensation in der Follow-up-Periode eine Accuracy von 0,84, eine AUC von 0,86 und einen F1-Score von 0,84 erreichten. Bezogen auf die Testdaten betrugen die Accuracy 0,73, die AUC 0,8 und der F1-Score 0,72.

Tabelle 20: Metriken neuronales Netz ungeclustert prospektiv

	Trainingsdaten	Testdaten
Accuracy	0,838	0,725
AUC	0,856	0,796
F1	0,838	0,724

Bezogen auf die prospektive Vorhersage durch neuronale Netze wurde der maximale Bilirubinwert am höchsten gerankt, gefolgt von den Baselinewerten von Natrium, PTT und Albumin (Abbildung 17)

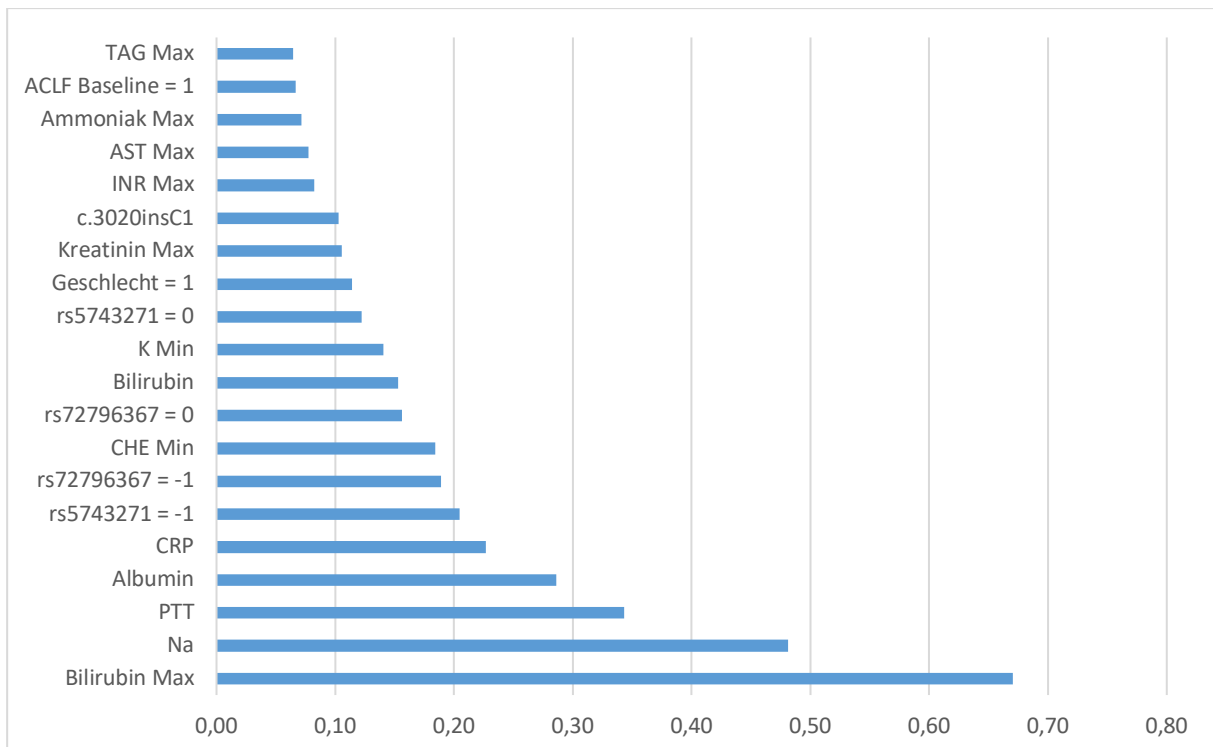


Abbildung 17: Top 20 neuronales Netz ungeclustert prospektiv

Gelclustert:

Tabelle 21 zeigt die Metriken für das gelclusterte Datenset. Auf den Trainingsdaten betrugen die Accuracy 0,82, die AUC 0,85 und der F1-Score 0,82, auf den Testdaten 0,73, 0,8 und 0,73.

Tabelle 21: Metriken neuronales Netz gelclustert prospektiv

	Trainingsdaten	Testdaten
Accuracy	0,823	0,730
AUC	0,850	0,801
F1	0,818	0,725

Auf dem gelclusterten Datenset wurde ebenfalls der maximale Bilirubinwert am höchsten gewichtet, es folgen die Baselinewerte von Natrium und Albumin sowie fehlende Informationen über die *NOD2*-Genotypen rs72796367 und rs5743271 (Abbildung 18).

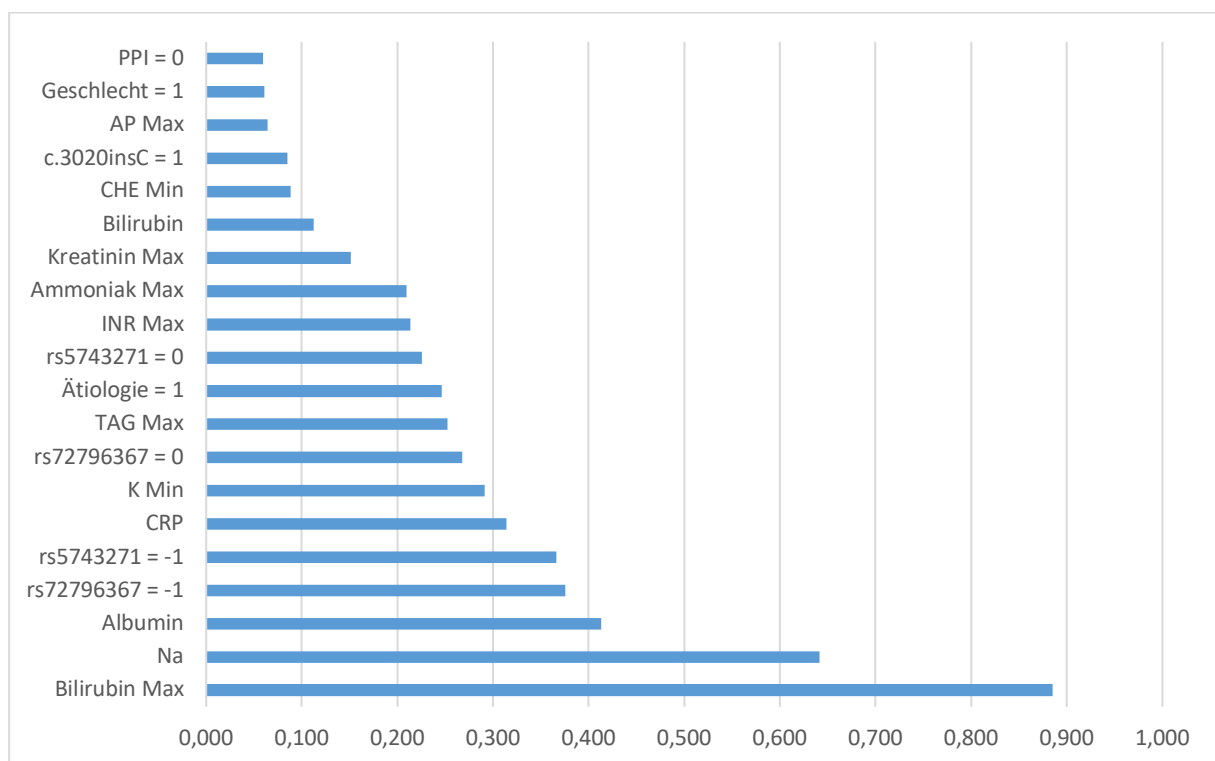


Abbildung 18: Top 20 neuronales Netz gelclustert prospektiv

4.5 Vergleich der Ergebnisse

Zur besseren Übersicht werden die jeweils zwanzig wichtigsten Parameter noch einmal vergleichend aufgelistet:

Retrospektiv ungeclustert:

Tabelle 22: Vergleich Top 20 retrospektiv ungeclustert

	Entscheidungsbaum	Random Forest	SVM	Neuronales Netz
1	Albumin	Bilirubin Max	Bilirubin	Bilirubin
2	Bilirubin Max	Albumin	Albumin	Albumin
3	rs5743271 = -1	Bilirubin	Na Min	Na Min
4	AFP Max	CRP	Albumin Min	Albumin Min
5	rs72796367 = -1	Varizen = 0	AST	rs72796367 = -1
6	TAG Max	AFP Max	rs72796367 = -1	Kreatinin
7	Bilirubin	Hb	rs5743271 = -1	Ferritin Min
8	CRP	Varizen = 1	Ferritin Min	rs5743271 = -1
9	PTT	rs5743271 = -1	rs5743271 = 0	Bilirubin Max
10	Ätiologie = 1	ALT Max	Ätiologie = 1	Langfristige Einnahme ATB = 1
11	Na Min	rs72796367 = -1	Kreatinin	ALT Max
12	Harnstoff	Ca	rs72796367 = 0	Varizen = 1
13	CHE Min	Na Min	Leukozyten Min	AST
14	Albumin Min	Harnstoff	CRP Max	rs72796367 = 0
15	Ammoniak Max	INR	K Max	Ca
16	AST	Ätiologie = 1	Varizen = 1	AP Max
17	rs5743271 = 0	Ferritin Min	Varizen = 0	rs5743271 = 0
18	ALT	AP Max	Ca	Leukozyten Min
19	ALT Max	AST	HB Min	Varizen = 0
20	Leukozyten	GGT Max	Bilirubin Max	Hb

Retrospektiv geclustert:**Tabelle 23:** Vergleich Top 20 retrospektiv geclustert

	Entscheidungsbaum	Random Forest	SVM	Neuronales Netz
1	Albumin	Albumin	Bilirubin	Bilirubin
2	Bilirubin Max	Bilirubin Max	Albumin	Albumin
3	AFP Max	Bilirubin	Na Min	Na Min
4	rs5743271 = -1	CRP	AST	rs5743271 = -1
5	rs72796367 = -1	AFP Max	rs5743271 = -1	rs72796367 = -1
6	Bilirubin	Hb	rs72796367 = -1	Bilirubin Max
7	CRP	Varizen = 0	rs5743271 = 0	Albumin Min
8	Albumin Min	ALT Max	Ferritin Min	Ferritin Min
9	TAG Max	rs72796367 = -1	ALT Max	ALT Max
10	CRP Max	Varizen = 1	Kreatinin	Kreatinin
11	Na Min	rs5743271 = -1	rs72796367 = 0	rs5743271 = 0
12	Ätiologie = 1	INR	CRP Max	Varizen = 1
13	ALT	ALT	Varizen = 1	AST
14	INR	Ätiologie = 1	Ätiologie = 1	Langfristige Einnahme ATB = 1
15	Ammoniak Max	Kreatinin	Albumin Min	rs72796367 = 0
16	Varizen = 0	Na Min	Varizen = 0	Hb
17	Hb	Langfristige Einnahme ATB = 0	Hb	Varizen = 0
18	AP Max	Langfristige Einnahme ATB = 1	Bilirubin Max	AP Max
19	Varizen = 1	rs72796367 = 0	Leukozyten Min	K Max
20	Ferritin Min	Ferritin Min	Ammoniak Max	K Min

Prospektiv ungeclustert

Tabelle 24: Vergleich Top 20 prospektiv ungeclustert

	Entscheidungsbaum	Random Forest	SVM	Neuronales Netz
1	Bilirubin Max	Bilirubin Max	Bilirubin Max	Bilirubin Max
2	Harnstoff Max	Albumin	Na	Na
3	Albumin	Kreatinin Max	Albumin	PTT
4	CHE Min	Bilirubin	PTT	Albumin
5	Kreatinin Max	CRP	Ammoniak Max	CRP
6	Ammoniak Max	Harnstoff Max	CRP	rs5743271 = -1
7	INR	PTT	rs5743271 = -1	rs72796367 = -1
8	Bilirubin	Ammoniak Max	rs72796367 = -1	CHE Min
9	CRP	PTT Max	rs72796367 = 0	rs72796367 = 0
10	CRP Max	Albumin Min	Harnstoff Max	Bilirubin
11	PTT	ALT Max	CHE Min	K Min
12	Albumin Min	INR Max	rs5743271 = 0	rs5743271 = 0
13	ALT Max	Hb	Kreatinin Max	Geschlecht = 1
14	Harnstoff	INR	Albumin Min	Kreatinin Max
15	Ferritin Min	CHE Min	PTT Max	c.3020insC = 1
16	Varizen = 0	Varizen = 1	K Min	INR Max
17	HB Min	TAG Max	AP Max	AST Max
18	Ca Min	K Min	Varizen = 1	Ammoniak Max
19	Na Min	Na	Ätiologie = 1	ACLF Baseline = 1
20	Ätiologie = 1	Ca Min	INR Max	TAG Max

Prospektiv geclustert

Tabelle 25: Vergleich Top 20 prospektiv geclustert

	Entscheidungsbaum	Random Forest	SVM	Neuronales Netz
1	Bilirubin Max	Bilirubin Max	Bilirubin Max	Bilirubin Max
2	Kreatinin Max	Kreatinin Max	Na	Na
3	INR	Albumin	Albumin	Albumin
4	CHE Min	CRP	CRP	rs72796367 = -1
5	Albumin	Bilirubin	Ammoniak Max	rs5743271 = -1
6	Albumin Min	Albumin Min	INR Max	CRP
7	Ammoniak Max	Ammoniak Max	Kreatinin Max	K Min
8	Varizen = 1	INR	CHE Min	rs72796367 = 0
9	CRP	Varizen = 1	Ätiologie = 1	TAG Max
10	Kreatinin	INR Max	INR	Ätiologie = 1
11	Bilirubin	CHE Min	rs5743271 = -1	rs5743271 = 0
12	CRP Max	Na	rs72796367 = -1	INR Max
13	Glucose Max	Varizen = 0	rs72796367 = 0	Ammoniak Max
14	Varizen = 0	AP Max	rs5743271 = 0	Kreatinin Max
15	AFP Max	K Min	Hb	Bilirubin
16	Ferritin Min	Ferritin Min	Albumin Min	CHE Min
17	TAG Max	Ätiologie = 1	TAG Max	c.3020insC = 1
18	Thrombozyten	Kreatinin	PPI = 1	AP Max
19	Leukozyten Min	Glucose Max	Geschlecht = 1	Geschlecht = 1
20	Thrombozyten Min	Hb	Ferritin Min	PPI = 0

5 Diskussion

5.1 Zusammenfassung der Ergebnisse

Im Rahmen dieser Arbeit wurde maschinelles Lernen zum ersten Mal zur Identifikation von Parametern, die in einer unselektierten Gruppe von Patienten mit Zirrhose eine Dekompensation vorhersagen, verwendet.

Für die Vorhersage einer früheren Dekompensation waren für Entscheidungsbäume und Random Forests der Baselinewert des Albumins und der maximale Bilirubinwert die höchstgewichteten Parameter, gefolgt vom Maximalwert des AFP bzw. vom Bilirubin-Baselinerwert. SVM und neuronale Netze gewichteten mit den Baselinewerten von Bilirubin und Albumin und dem maximalen Bilirubinwert die gleichen Parameter am höchsten.

Für die Vorhersage einer Dekompensation in der Follow-up-Periode war für alle Verfahren der maximale Bilirubinwert der am höchsten gewichtete Parameter. Für Entscheidungsbäume und Random Forests folgen der maximale Kreatininwert und der Baselinewert der INR bzw. des Albumins, für SVM und neuronale Netze die Baselinewerte von Natrium und Albumin.

5.2 Vergleich der Ergebnisse der einzelnen Verfahren

Random Forests und SVM performten am besten, neuronale Netze und Entscheidungsbäume lieferten ebenfalls gute Ergebnisse. Bezüglich der Performancemetriken unterscheiden sich die ungeclusterten Modelle kaum von den geclusterten Modellen, die höchstgewichteten Parameter basierend auf der PFI sind insgesamt ebenfalls vergleichbar. Bei der ungeclusterten Analyse fällt auf, dass bei der prospektiven Analyse der (maximale) Harnstoffspiegel sowie die PTT hoch gewichtet wurden.

Auch wenn sich die Top 20 Parameter der einzelnen Verfahren in Details bzw. in der Reihenfolge etwas unterscheiden, sind die Ergebnisse insgesamt doch vergleichbar.

Entscheidungsbäume haben den Vorteil, dass sie einfach und gut verständlich sind sowie die geringste Rechenzeit benötigen. Allerdings neigen sie sehr stark zu Overfitting, was auch hier deutlich wurde. Random Forests sind ebenfalls relativ einfach, sie zeigten eine sehr gute Performance. SVM sind vom Prinzip her etwas anspruchsvoller als Random Forests, sie erreichten in der prospektiven Analyse die besten Ergebnisse. Ein Vorteil von SVM liegt darin, dass sie auch in höherdimensionalen Räumen effektiv sind und auch auf vergleichsweise kleinen Trainingsmengen arbeiten können. Neuronale Netze

werden v. a. in der Bilderkennung/-analyse eingesetzt, zeigten hier aber auch eine gute Leistung. Bei ihnen spielt das Problem von Overfitting aber eine größere Rolle als bei Random Forests oder SVM.

5.3 Diskussion der Parameter

Die höchstgewichteten Variablen lassen sich grob in verschiedene Gruppen einteilen, u. a. in klassische Lebersynthese- und Cholestaseparameter, Parameter, die mit einer portalen Hypertension assoziiert sind und infektassoziierte Parameter.

So spielte der Bilirubinspiegel bzw. dessen Maximalwert als klinisch überaus relevanter Laborwert für die Einschätzung der Leberfunktion auch bei allen hier getesteten Verfahren sowohl in der retrospektiven als auch in der prospektiven Analyse eine entscheidende Rolle.

Die klassischen Lebersyntheseparameter wie Albumin und INR-Wert hatten ebenfalls einen großen Einfluss auf die Modelle. Bei der prospektiven Vorhersage befand sich zusätzlich die minimale Cholinesterase als Syntheseparameter unter den Top 20 Parametern. Im ungeclusterten Modell wurde zudem die PTT hoch gewichtet, auch diese verändert sich bei eingeschränkter Synthesekapazität der Leber.

Variablen, die mit einer portalen Hypertension assoziiert werden, beeinflussten die Vorhersage einer Dekompensation ebenfalls stark. Passend dazu wurde bei der prospektiven Vorhersage durch SVM und neuronale Netze auch der Natriumwert sehr hoch gerankt. Dies kann man als Hinweis auf eine Hypervolämie bzw. verstärkte Flüssigkeitsretention mit konsekutiver Verdünnungshyponatriämie als wichtigen pathophysiologischen Schritt bei der Ausbildung einer portalen Hypertension bzw. deren Komplikationen interpretieren. Ösophagusvarizen, die direkt mit einer portalen Hypertension assoziiert sind, befanden sich außer bei der prospektiven Vorhersage mit SVM und neuronalen Netzen auch unter den Top 20 Parametern. Überraschenderweise gehörte die Thrombozytenzahl, die ebenfalls sehr häufig als Marker für eine portale Hypertension genutzt wird, jedoch nur bei Entscheidungsbäumen in der prospektiven Analyse zu den wichtigsten Parametern basierend auf der PFI.

Auffällig ist, dass Parameter, die mit Infektionen assoziiert sind, eine wichtige Rolle bei der Vorhersage einer Dekompensation spielten. So befanden sich *NOD2*-Genotypen, CRP-Werte und Informationen über eine langfristige antibiotische Prophylaxe unter den 20 höchstgewichteten Parametern. Passend dazu sind Infektionen schon seit längerem als Auslöser einer Dekompensation bekannt. Dass insbesondere auch fehlende Informationen über *NOD2*-Genotypen hoch gewichtet wurden, kann möglicherweise dadurch erklärt werden, dass klinisch schwerer erkrankte Patienten häufiger genotypisiert wurden.

Nierenwerte wie das (maximale) Kreatinin sowie der (maximale) Harnstoffspiegel waren als Marker für die Entwicklung eines hepatorenenalen Syndroms insbesondere in der prospektiven Vorhersage ebenfalls relevante Parameter.

Der hohe Einfluss der Harnstoffwerte im ungeclusterten Modell könnte auch zusätzlich als Hinweis auf eine hepatische Enzephalopathie gedeutet werden. Passend dazu befand sich der maximale Ammoniakwert, der eindeutig mit einer HE assoziiert wird, in der prospektiven Analyse bei allen Verfahren der unter den wichtigsten Parametern.

Ebenfalls in der retrospektiven Analyse wurden die Transaminasen bzw. dessen Maximalwerte als Maß für eine Leberzellschädigung als relevante Werte betrachtet.

Überraschenderweise befand sich der minimale Ferritinwert bei den meisten Modellen unter den Top 20 Parametern, dies kann durch einen Eisenverlust im Rahmen von (chronischen) Blutungen oder durch eine Mangelernährung erklärt werden.

5.4 Aktuelle Studienlage bezüglich Dekompensation

Studien haben gezeigt, dass der HVPG als Marker der portalen Hypertension Dekompensationen vorhersagt [98]. Da die Messung des HVPG jedoch ein invasives Verfahren darstellt, ist seine Verwendung als Screening-Parameter schwierig, insbesondere bei kompensierten Patienten.

Gemäß dem Baveno-VII-Konsens kann eine klinisch signifikante portale Hypertension nun auch auf nicht-invasive Weise mittels Lebersteifigkeitsmessung (LSM) +/- Thrombozytenzahl diagnostiziert werden [36]. Darüber hinaus haben Reiniš et al. gezeigt, dass der Schweregrad der portalen Hypertension durch maschinelles Lernen bewertet werden kann. Dazu testeten sie verschiedene ML-Methoden und bestimmten die drei bzw. fünf wichtigsten Parameter durch eine rekursive Featureelimination. Sie fanden heraus, dass ein 3-Parameter-Modell mit Thrombozytenzahl, Bilirubin und INR oder ein 5-Parameter-Modell mit Thrombozytenzahl, Bilirubin, aktivierter partieller Thromboplastinzeit, GGT und CHE in Kombination mit LSM die besten Ergebnisse erzielte [96].

Wie bereits beschrieben, wurden auch in der hier vorliegenden Arbeit Variablen, die mit einer portalen Hypertension assoziiert sind, hoch gewichtet.

Infektionen sind schon seit längerem als Trigger einer Dekompensation bekannt. Bothou et al. konnten bereits Entzündungsmarker wie IL-6 als Prädiktoren für eine Dekompensation identifizieren [18]. Dies wird durch das hier vorliegende Ergebnis unterstützt, dass der *NOD2*-Genotyp mit den Varianten p.R702W, p.G908R und c.3020insC, die nicht nur mit frühem Tod, sondern auch mit SBP und anderen

Infektionen in Verbindung gebracht werden, ebenfalls als wichtiger Faktor basierend auf der PFI angesehen wurde. Interessanterweise wurden die Varianten rs5743271 und rs72796367 - deren prognostischer Wert bei Leberzirrhose bisher nicht untersucht wurde - auf Grundlage der PFI höher eingestuft als die besser charakterisierten Varianten p.R702W, p.G908R und c.3020insC. Allerdings fehlten für eine relevante Anzahl von Patienten Daten für die Genotypen rs5743271 und rs72796367, was die Validität des Modells einschränkt. Künftige unabhängige Studien müssen die prognostische Bedeutung dieser Varianten bei Zirrhose überprüfen.

In weiteren Studien wurden Anämie und erniedrigte Vitamin-D-Spiegel als Prädiktoren für eine Dekompensation ermittelt [18,73]. Der Hämoglobinwert befand sich in der retrospektiven Analyse bei allen Verfahren unter den 20 wichtigsten Parametern, bei der prospektiven Vorhersage nur bei den Random Forests und SVM. Der Vitamin-D-Spiegel war nicht in der Datenbank eingeschlossen, sodass darüber keine Aussage getroffen werden kann.

Der 2022 neu entwickelte EPOD-Score [104] ist der einzige Score, der sich auf Dekompensation und nicht auf Mortalität bezieht. Er verwendet die Thrombozytenzahl sowie die Albumin- und Bilirubinkonzentration, um Dekompensationen in einem 3-Jahres-Zeitraum vorherzusagen. Verglichen zum EPOD-Score ergab diese Studie ebenfalls eine hohe Bedeutung von Albumin und Bilirubin, während die Thrombozytenzahl wie bereits beschrieben überraschend niedrig gewichtet wurde.

5.5 Vergleich der Ergebnisse mit bestehenden Scores

Die in dieser Studie identifizierten Parameter stimmen zumindest teilweise mit den Parametern überein, die in etablierten Scores wie Child-Pugh, MELD, ALBI und Fib-4 verwendet werden.

Mit dem Child-Pugh-Score [91] werden Patienten mit Zirrhose anhand von Gesamtbilirubin, Serumalbumin und INR-Wert sowie dem Auftreten von Aszites und hepatischer Enzephalopathie in drei Stadien (Child A-C) eingeteilt, die eine prognostische Aussagekraft haben. Albumin und Bilirubin gehörten bei allen getesteten Verfahren auch zu den relevantesten Parametern für die Einschätzung einer Dekompensation, der INR-Wert war v. a. in der prospektiven Beurteilung bedeutsam. Da Aszites und eine hepatische Enzephalopathie Dekompensationskriterien darstellen, sind diese Parameter zur Prognose einer Dekompensation hinfällig.

Der MELD-Score [63,64] ist ein bewährtes Modell zur Vorhersage der Überlebenswahrscheinlichkeit von Patienten mit Lebererkrankungen im Endstadium.

Im Laufe der Zeit wurden verschiedene Modifikationen des MELD-Scores vorgeschlagen, aber die Werte von Bilirubin, Kreatinin und INR wurden stets berücksichtigt. Ursprünglich wurde auch die

Ätiologie der Erkrankung im MELD-Score betrachtet. Es hat sich jedoch gezeigt, dass sie für die Bestimmung des Überlebens weniger wichtig ist, weshalb sie aus dem Modell entfernt wurde. Passend dazu war in dieser Analyse auch nur eine alkoholische Ätiologie relevant. Mehrere Studien haben gezeigt, dass die Einbeziehung von Natrium in den MELD-Score dessen Vorhersagegenauigkeit erhöht [13,69]. In Übereinstimmung damit wurde der Serumnatriumwert durch die PFI insbesondere bei der prospektiven Vorhersage als wichtiger Parameter eingeschätzt.

Darüber hinaus wurde der MELD-Plus-Score [66] als Verbesserung des MELD(-Na)-Scores vorgeschlagen. Neben den Parametern des MELD-Na-Scores berücksichtigt er Serumalbumin, Leukozytenzahl, Serumcholesterin, Alter und die Dauer des Krankenhausaufenthalts. Dementsprechend wurden auch die Werte von Gesamtbilirubin, Serumalbumin, Kreatinin und Natrium durch die PFI hoch eingestuft. Der INR-Wert wurde insbesondere für die Vorhersage einer Dekompensation in der Nachbeobachtungszeit ebenfalls hoch gewichtet, während der Gesamtcholesterinspiegel nicht zu den 20 wichtigsten Parametern gehörte. Da im Rahmen der Clusteranalyse der CRP-Wert anstelle der Leukozytenzahl als Entzündungsmarker ausgewählt wurde, kann die Relevanz der Leukozytenzahl hier nicht direkt beurteilt werden, aber der CRP-Wert war eindeutig ein wichtiger Parameter. Da nur Parameter berücksichtigt werden sollten, die in einem pathophysiologischen/kausalen Zusammenhang mit der Krankheit stehen könnten, wurde das Patientenalter in dieser Analyse nicht berücksichtigt.

Der ALBI-Score [53,62,86,89,111] kombiniert Serumalbumin und Gesamtbilirubin zur Bewertung der Leberfunktion; diese Parameter wurden auch durch die PFI bei allen getesteten Verfahren als relevant eingestuft.

Der Fib-4-Index [109] schätzt das Ausmaß der Fibrose anhand von Alter, AST, ALT und Thrombozytenzahl. Er ist umfassend validiert und hat eine prognostische Aussagekraft für Patienten mit chronischen Lebererkrankungen. Die Bedeutung der Thrombozytenzahl war in dieser Studie geringer als erwartet, die Serum-Aminotransferase-Aktivitäten befanden sich lediglich in der retrospektiven Analyse unter den Top 20 Parametern.

Der CLIF-C AD-Score [60] prognostiziert die Sterblichkeit hospitalisierter Patienten mit Zirrhose und akuter Dekompensation basierend auf Alter, Leukozyten, Kreatinin, INR und Natrium. Wie bereits beschrieben, wurden diese Parameter auch durch die PFI überwiegend hoch eingestuft.

Die Ergebnisse stimmen nur teilweise mit denen von Guo et al. [51] überein, die einen Datensatz mit Daten von 34.575 Patienten mithilfe neuronaler Netze, Random Forests und logistischer Regression analysierten, um die 30-, 90- und 180-Tage-Sterblichkeit bei Patienten mit Zirrhose vorherzusagen. Sie zeigten, dass alle Modelle den MELD-Score übertrafen, wobei neuronale Netze am besten abschnitten. Zusätzlich zu den Variablen des MELD-Scores waren AP, ALT und Hämoglobin die wichtigen Parameter in ihrer Analyse. In dieser Analyse konnte die AP nicht als wichtiger Parameter bestätigt werden, der Hämoglobinwert befand sich in der retrospektiven Analyse bei allen Verfahren und bei der

prospektiven Vorhersage bei den baumbasierten Verfahren unter den 20 wichtigsten Parametern basierend auf der PFI.

Sowohl Lin et al. als auch Yu et al. testeten ebenfalls verschiedene ML-Techniken, um die Mortalität von Patienten mit Lebererkrankungen im Endstadium abzuschätzen. Yu et al. [120] konnten durch eine klassische multivariate logistische Regression und Cox-Regression neben dem INR-Wert den Harnstoffspiegel sowie CRP-Wert als signifikante Parameter identifizieren. Bei den maschinellen Lernverfahren wurden Harnstoff und Alter als wichtigste Parameter betrachtet, gefolgt von Kreatinin, INR-Wert und Bilirubin. In der Studie von Lin et al. [76] waren ebenfalls der Harnstoffwert, sowie Bilirubin und Natrium die entscheidenden Parameter. Mithilfe unüberwachter Lernmodelle konnten sie zudem Kreatinin, Hämoglobin und Albumin als signifikante Parameter bestimmen. In der hier vorliegenden Studie wurde zwar im Rahmen der Clusteranalyse Kreatinin statt Harnstoff als Variable selektiert, in der ungelusterten Analyse wurde der (maximale) Harnstoffwert jedoch insbesondere bei der prospektiven Vorhersage einer Dekompensation hoch gewichtet.

Kanwal et al. nutzen ebenfalls ein einfaches ML-Modell, um die Mortalität von Patienten mit Zirrhose vorherzusagen. In ihrer Studie waren Alter, Bilirubin, Albumin, Hb, Natrium, Thrombozytenzahl, eine AST:ALT-Ratio > 2 , sowie das Vorliegen von Aszites, einer hepatischen Enzephalopathie oder eines HCC sowie „black race“ signifikante Parameter [65].

Wie bereits beschrieben, bezieht sich jedoch keines der in diesem Abschnitt genannten Modelle auf Dekompensationen, sondern i. d. R. auf die Mortalität. Daher ist ein direkter Vergleich mit den Ergebnissen dieser Studie schwierig.

5.6 Einsatz von maschinellem Lernen

Maschinelle Lernverfahren haben im Vergleich zu einer klassischen statistischen Analyse einige Vorteile: Sie erlauben eine nichtlineare Klassifikation und sind leicht skalierbar, können also auch auf großen Datenmengen mit vielen Variablen trainiert werden. In vielen Studien performten Modelle, die mit maschinellem Lernen entwickelt wurden, besser als die auf „klassischer“ Statistik basierenden Modelle. Durch eine ständige Aktualisierung der Datenbasis können die Modelle laufend weitertrainiert und damit verbessert werden.

Eine akkurate Vorhersage des individuellen Krankheitsverlaufes könnte direkte Konsequenzen für das klinische Management von Patienten haben. Eine genaue Vorhersage würde es den Ärzten beispielsweise erlauben, den Patienten die Dringlichkeit von Lebensstilveränderungen nahezubringen

und die pharmakologische Therapie sowie die Kontrollintervalle zielgerichtet an das individuelle Risiko anzupassen [81].

Allerdings besteht eine große Diskrepanz zwischen der enormen Anzahl an wissenschaftlichen Publikationen, die sich mit Machine Learning bzw. Deep Learning im Bereich der Hepatologie befassen, und dem sehr limitierten klinischen Einsatz. Dies könnte u. a. darin begründet sein, dass bei den meisten ML-Studien eine externe Validierung fehlt und die Studien oft retrospektiv angelegt sind. In der Klinik fehlt auch häufig die benötigte technische Infrastruktur, bei einem großen Teil des Personals fehlt Wissen und Vertrauen in die ML-Verfahren und es treten immer wieder Datenschutzbedenken auf [2,70]. Oft wird eine mangelnde Transparenz der Entscheidungen bemängelt, zudem sind immer noch einige rechtliche sowie ethische Fragen ungeklärt. Nicht zuletzt können mit ML immer nur Korrelationen, aber keine Kausalitäten nachgewiesen werden.

5.7 Stärken der Studie

Bei der verwendeten Studie handelt es sich um eine multizentrische Studie, die an zwei deutschen Universitätskliniken durchgeführt wurde und mit insgesamt 983 Patienten eine große Datenbasis aufweist. Zumindest ein Teil der Studie hatte ein prospektives Design. Es wurde eine Vielzahl von Parametern berücksichtigt, neben klinischen Angaben und Laborwerten wurden auch genetische Daten, insbesondere der *NOD2*-Genotyp betrachtet. In anderen Studien im Bereich der Hepatologie, die mit maschinellem Lernen arbeiten, werden nur selten die von dem jeweiligen Modell genutzten Parameter identifiziert und genetische Daten werden ebenfalls nur unzureichend untersucht.

Durch das maschinelle Lernen wurde eine vielversprechende neue Methode für die Analyse großer Datenmengen angewendet, die bei verwandten Fragestellungen bereits erfolgreich eingesetzt wird.

5.8 Limitationen der Studie

Die Studie weist jedoch auch eine Reihe von Limitationen auf.

Zunächst einmal ist die Anwendung von maschinellen Lernverfahren mit mehreren Einschränkungen verbunden. Die meisten maschinellen Lernverfahren können nicht mit fehlenden Werten umgehen, aber Patientendaten sind oft unvollständig und schwer zu beschaffen. In der INCA-Studiendatenbank fehlte aufgrund des Screening-Charakters der Datenerhebung ein großer Teil der Daten für einige Parameter.

Diese mussten für die Analyse imputiert werden, was die Ergebnisse verfälscht haben könnte. Obwohl es mehrere Ansätze gibt, um die Verfahren erklärbarer zu machen, ist es immer noch schwierig, Parameter mit guter Vorhersagekraft zu ermitteln. Die PFI ist ein Ansatz, um die Bedeutung von Merkmalen zu bestimmen, wird aber durch korrelierte Merkmale stark verzerrt.

Zweitens fehlt der Studie eine unabhängige externe Validierungskohorte. Obwohl ein Holdout-Testdatensatz verwendet wurde, ist eine weitere externe Validierung erforderlich. Außerdem war ein Teil der Studie retrospektiv angelegt und für den prospektiven Teil war die mittlere Nachbeobachtungszeit mit 12 Monaten eher kurz. Darüber hinaus war der Umfang des Datensatzes relativ klein, insbesondere für den prospektiven Teil, und alle Patienten wurden an tertiären Referenzzentren rekrutiert, was die Verallgemeinerbarkeit auf andere Gesundheitseinrichtungen ebenfalls einschränken könnte. Außerdem enthielt die Datenbank keine Daten zur Hospitalisierung und bei den Patienten aus Halle wurden die Laborwerte nur zum Zeitpunkt der Aufnahme erhoben. Daten zur Virusreplikation oder zur antiviralen Therapie bei Patienten mit viraler Hepatitis oder zur Abstinenz von Patienten mit alkoholbedingten Erkrankungen lagen nicht vor. Außerdem war ein Drittel der Patienten zum Zeitpunkt der Aufnahme in die Studie dekompenziert, was die Ergebnisse der PFI-Analyse verfälscht haben könnte, da die aktuelle Dekompensation mit abnormen Leberwerten verbunden ist. Überraschenderweise gehörte die Thrombozytenzahl nur in der prospektiven Analyse zu den 20 wichtigsten Parametern. Ein direkter statistischer Vergleich des Modells mit etablierten Scores war aufgrund der unterschiedlichen Endpunkte nicht möglich.

5.9 Zusammenfassung und Ausblick

Trotz der Einschränkungen der Studie wurde eine vielversprechende neue Methode angewandt, um Risikofaktoren bzw. Marker zu ermitteln, die das Risiko einer Dekompensation beeinflussen könnten. Unter den getesteten ML-Verfahren erzielten SVM und Random Forests die besten Ergebnisse, aber auch neuronale Netze und Entscheidungsbäume zeigten gute Ergebnisse.

Zusätzlich zu Bilirubinspiegel und Lebersyntheseparametern wie Albumin und INR spielten mit einer portalen Hypertension assoziierte Faktoren eine wichtige Rolle. Außerdem waren *NOD2*-Genotypen und Parameter, die mit Infektionen bzw. Entzündungen in Verbindung gebracht werden (wie etwa CRP-Werte und eine langfristige Antibiotikaeinnahme), von hoher Relevanz.

Weitere Forschungsarbeiten sind erforderlich, um Risikomarker und Risikofaktoren zu unterscheiden, damit die Prävention der Komplikationen bei portaler Hypertonie und Zirrhose optimiert werden kann.

6 Literaturverzeichnis

1. Abraldes JG, Iwakiri Y, Loureiro-Silva M, Haq O, Sessa WC, Groszmann RJ (2006) Mild increases in portal pressure upregulate vascular endothelial growth factor and endothelial nitric oxide synthase in the intestinal microcirculatory bed, leading to a hyperdynamic state. *Am J Physiol Gastrointest Liver Physiol* 290:G980-987
2. Agrawal R, Prabakaran S (2020) Big data in digital healthcare: lessons learnt and recommendations for general practice. *Heredity* 124:525–534
3. Anders U (1997) Statistische neuronale Netze. München: Vahlen
4. Appenrodt B, Grünhage F, Gentemann MG, Thyssen L, Sauerbruch T, Lammert F (2010) Nucleotide-binding oligomerization domain containing 2 (*NOD2*) variants are genetic risk factors for death and spontaneous bacterial peritonitis in liver cirrhosis. *Hepatology* 51:1327–1333
5. Arvaniti V, D’Amico G, Fede G, Manousou P, Tsochatzis E, Pleguezuelo M, Burroughs AK (2010) Infections in patients with cirrhosis increase mortality four-fold and should be used in determining prognosis. *Gastroenterology* 139:1246–1256, 1256.e1–5
6. Backhaus K, Erichson B, Weiber R (2015) Fortgeschrittene multivariate Analysemethoden: eine anwendungsorientierte Einführung. Springer-Verlag
7. Bajaj JS, Zadornova Y, Heuman DM, Hafeezullah M, Hoffmann RG, Sanyal AJ, Saeian K (2009) Association of proton pump inhibitor therapy with spontaneous bacterial peritonitis in cirrhotic patients with ascites. *Am J Gastroenterol* 104:1130–1134
8. Banerjee R, Das A, Ghoshal UC, Sinha M (2003) Predicting mortality in patients with cirrhosis of liver with application of neural network technology. *J Gastroenterol Hepatol* 18:1054–1060
9. Bergstra J, Bengio Y (2012) Random search for hyper-parameter optimization. *J Mach Learn Res* 13:281–305
10. Bernardi M, Gitto S, Biselli M (2011) The MELD score in patients awaiting liver transplant: strengths and weaknesses. *J Hepatol* 54:1297–1306
11. Bertsimas D, Kung J, Trichakis N, Wang Y, Hirose R, Vagefi PA (2019) Development and validation of an optimized prediction of mortality for candidates awaiting liver transplantation. *Am J Transplant Off J Am Soc Transplant Am Soc Transpl Surg* 19:1109–1118
12. Berzigotti A, Seijo S, Reverter E, Bosch J (2013) Assessing portal hypertension in liver diseases. *Expert Rev Gastroenterol Hepatol* 7:141–155
13. Biggins SW, Kim WR, Terrault NA, Saab S, Balan V, Schiano T, Benson J, Therneau T, Kremers W, Wiesner R, Kamath P, Klintmalm G (2006) Evidence-based incorporation of serum sodium concentration into MELD. *Gastroenterology* 130:1652–1660
14. Borzio M, Salerno F, Piantoni L, Cazzaniga M, Angeli P, Bissoli F, Boccia S, Colloredo-Mels G, Corigliano P, Fornaciari G, Marengo G, Pitarà R, Salvagnini M, Sangiovanni A (2001) Bacterial infection in patients with advanced cirrhosis: a multicentre prospective study. *Dig Liver Dis Off J Ital Soc Gastroenterol Ital Assoc Study Liver* 33:41–48
15. Bosch J (2010) Hepatic endothelial dysfunction and abnormal angiogenesis: New targets in the treatment of portal hypertension. *J Hepatol* 53:10

16. Bosch J, Chung C, Carrasco-Zevallos OM, Harrison SA, Abdelmalek MF, Shiffman ML, Rockey DC, Shanis Z, Juyal D, Pokkalla H, Le QH, Resnick M, Montalto M, Beck AH, Wapinski I, Han L, Jia C, Goodman Z, Afdhal N, Myers RP, Sanyal AJ (2021) A Machine Learning Approach to Liver Histological Evaluation Predicts Clinically Significant Portal Hypertension in NASH Cirrhosis. *Hepatology* 74:3146–3160
17. Boser BE, Guyon IM, Vapnik VN (1992) A training algorithm for optimal margin classifiers. Association for Computing Machinery, New York, NY, USA. URL: <https://doi.org/10.1145/130385.130401>
18. Bothou C, Rüschenbaum S, Kubesch A, Quenstedt L, Schwarzkopf K, Welsch C, Zeuzem S, Welzel TM, Lange CM (2020) Anemia and Systemic Inflammation Rather than Arterial Circulatory Dysfunction Predict Decompensation of Liver Cirrhosis. *J Clin Med* 9:1263
19. Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) Classification and regression trees. Routledge
20. Breiman L (2001) Random Forests. *Mach Learn* 45:5–32
21. Bruns T, Peter J, Reuken PA, Grabe DH, Schuldes SR, Brenmoehl J, Schölmerich J, Wiest R, Stallmach A (2012) *NOD2* gene variants are a risk factor for culture-positive spontaneous bacterial peritonitis and monomicrobial bacterascites in cirrhosis. *Liver Int* 32:223–230
22. Cárdenas A, Bataller R, Arroyo V (2000) Mechanisms of ascites formation. *Clin Liver Dis* 4:447–465
23. Casper M, Mengel M, Fuhrmann C, Herrmann E, Appenrodt B, Schiedermaier P, Reichert M, Bruns T, Engelmann C, Grünhage F, Lammert F, INCA trial group (2015) The INCA trial (Impact of *NOD2* genotype-guided antibiotic prevention on survival in patients with liver Cirrhosis and Ascites): study protocol for a randomized controlled trial. *Trials* 16:83
24. Chavez-Tapia NC, Barrientos-Gutierrez T, Tellez-Avila F, Soares-Weiser K, Mendez-Sanchez N, Gluud C, Uribe M (2011) Meta-analysis: antibiotic prophylaxis for cirrhotic patients with upper gastrointestinal bleeding - an updated Cochrane review. *Aliment Pharmacol Ther* 34:509–518
25. Chen R-C, Cai Y-J, Wu J-M, Wang X-D, Song M, Wang Y-Q, Zheng M-H, Chen Y-P, Lin Z, Shi KQ (2017) Usefulness of albumin-bilirubin grade for evaluation of long-term prognosis for hepatitis B-related cirrhosis. *J Viral Hepat* 24:238–245
26. Choi KJ, Jang JK, Lee SS, Sung YS, Shim WH, Kim HS, Yun J, Choi J-Y, Lee Y, Kang B-K, Kim JH, Kim SY, Yu ES (2018) Development and Validation of a Deep Learning System for Staging Liver Fibrosis by Using Contrast Agent-enhanced CT Images in the Liver. *Radiology* 289:688–697
27. Cholongitas E, Senzolo M, Triantos C, Samonakis D, Patch D, Burroughs AK (2005) MELD is not enough--enough of MELD? *J Hepatol* 42:475–477; author reply 478-479
28. Cholongitas E, Papatheodoridis GV, Vangeli M, Terreni N, Patch D, Burroughs AK (2005) Systematic review: The model for end-stage liver disease--should it replace Child-Pugh's classification for assessing prognosis in cirrhosis? *Aliment Pharmacol Ther* 22:1079–1089
29. Chu CM, Chang KY, Liaw YF (1995) Prevalence and prognostic significance of bacterascites in cirrhosis with ascites. *Dig Dis Sci* 40:561–565

30. Cucchetti A, Vivarelli M, Heaton ND, Phillips S, Piscaglia F, Bolondi L, La Barba G, Foxton MR, Rela M, O'Grady J, Pinna AD (2007) Artificial neural network is superior to MELD in predicting mortality of patients with end-stage liver disease. *Gut* 56:253–258
31. D'Amico G (2001) Natural history of compensated cirrhosis and varices. *Complicat Cirrhosis Pathog Consequences Ther VA Am Assoc Study Liver Dis* 118–123
32. D'Amico G, Garcia-Tsao G, Pagliaro L (2006) Natural history and prognostic indicators of survival in cirrhosis: a systematic review of 118 studies. *J Hepatol* 44:217–231
33. D'Amico G, Pasta L, Morabito A, D'Amico M, Caltagirone M, Malizia G, Tinè F, Giannuoli G, Traina M, Vizzini G, Politi F, Luca A, Virdone R, Licata A, Pagliaro L (2014) Competing risks and prognostic stages of cirrhosis: a 25-year inception cohort study of 494 patients. *Aliment Pharmacol Ther* 39:1180–1193
34. De Franchis R, Salerno F (2002) Pathogenesis of ascites and predictors of resistance to therapy. *J Gastroenterol Hepatol* 17 Suppl 3:S242-247
35. De Franchis R (2015) Expanding consensus in portal hypertension: Report of the Baveno VI Consensus Workshop: Stratifying risk and individualizing care for portal hypertension. *J Hepatol* 63:743–752
36. De Franchis R, Bosch J, Garcia-Tsao G, Reiberger T, Ripoll C, Abraldes JG, Albillos A, Baiges A, Bajaj J, Bañares R, Barrufet M, Benajiba L, Berzigotti A, Bureau C, Calvaruso V, Cardenas A, D'Amico G, De Gottardi A, Dell'Era A, Escorsell A, Fallowfield J, Ferral H, Francque S, Gaba R, Garcia-Pagàn JC, Genescà J, Rodrigues SG, Gracia-Sancho J, Han G, Hernandez-Gea V, Jia J, Kiladjian JJ, Krag A, Laleman W, La Mura V, Lens S, Luo X, Mandorfer M, Murad SD, Paradis V, Patch D, Piano S, Pinzani M, Plessier A, Primignani M, Procopet B, Rautou PE, Rudler M, Sarin SK, Schepis F, Senzolo M, Shah V, Shukla A, Tandon P, Tellez L, Thabut D, Thiele M, Trebicka J, Tripathi D, Tsochatzis E, Turco L, Turon F, Valla D, Villanueva C, Wanless I, Yoshiji H (2022) Baveno VII – Renewing consensus in portal hypertension. *J Hepatol* 76:959–974
37. Decharatanachart P, Chaiteerakij R, Tiyyarattanachai T, Treeprasertsuk S (2021) Application of artificial intelligence in chronic liver diseases: a systematic review and meta-analysis. *BMC Gastroenterol* 21:10
38. Dionigi E, Garcovich M, Borzio M, Leandro G, Majumdar A, Tsami A, Arvaniti V, Roccarina D, Pinzani M, Burroughs AK, O'Beirne J, Tsochatzis EA (2017) Bacterial Infections Change Natural History of Cirrhosis Irrespective of Liver Disease Severity. *Am J Gastroenterol* 112:588–596
39. Ferenci P, Lockwood A, Mullen K, Tarter R, Weissenborn K, Blei AT (2002) Hepatic encephalopathy—Definition, nomenclature, diagnosis, and quantification: Final report of the Working Party at the 11th World Congresses of Gastroenterology, Vienna, 1998. *Hepatology* 35:716–721
40. Fernández J, Navasa M, Gómez J, Colmenero J, Vila J, Arroyo V, Rodés J (2002) Bacterial infections in cirrhosis: Epidemiological changes with invasive procedures and norfloxacin prophylaxis. *Hepatology* 35:140–148
41. Fisher RA (1936) The use of multiple measurements in taxonomic problems. *Ann Eugen* 7:179–188
42. Foreman MG, Mannino DM, Moss M (2003) Cirrhosis as a Risk Factor for Sepsis and Death: Analysis of the National Hospital Discharge Survey. *Chest* 124:1016–1020

43. García-Pagán J-C, Gracia-Sancho J, Bosch J (2012) Functional aspects on the pathophysiology of portal hypertension in cirrhosis. *J Hepatol* 57:458–461
44. Garcia-Tsao G, Abraldes JG, Berzigotti A, Bosch J (2017) Portal hypertensive bleeding in cirrhosis: Risk stratification, diagnosis, and management: 2016 practice guidance by the American Association for the study of liver diseases. *Hepatology* 65:310–335
45. Gerbes AL, Labenz J, Appenrodt B, Dollinger M, Gundling F, Gülberg V, Holstege A, Lynen-Jansen P, Steib CJ, Trebicka J, Wiest R, Zipprich A, Collaborators:, Deutsche Gesellschaft für Allgemein- und Viszeralchirurgie (DGAV), Deutsche Gesellschaft für Innere Medizin e.V. (DGIM), Deutsche Gesellschaft für Infektiologie e.V. (DGI), Deutsche Gesellschaft für Pathologie e.V./Bundesverband deutscher Pathologen e.V. (DGP/BDP), Deutsche Röntgengesellschaft e.V. (DRG), Deutsche Gesellschaft für Interventionelle Radiologie und minimal-invasive Therapie (DeGIR), Deutsche Gesellschaft für Nephrologie (DGfN), Deutsche Gesellschaft für Ultraschall in der Medizin (DEGUM), Deutsche Gesellschaft für Neurologie e.V. (DGN), Deutsche Gesellschaft für Ernährungsmedizin e.V. (DGEM), Lebertransplantierte Deutschland e.V. (2019) Aktualisierte S2k-Leitlinie der Deutschen Gesellschaft für Gastroenterologie, Verdauungs- und Stoffwechselkrankheiten (DGVS) „Komplikationen der Leberzirrhose“: AWMF-Nr.: 021-017. *Z Für Gastroenterol* 57:611–680
46. Ginés P, Rimola A, Planas R, Vargas V, Marco F, Almela M, Forné M, Miranda ML, Llach J, Salmerón JM (1990) Norfloxacin prevents spontaneous bacterial peritonitis recurrence in cirrhosis: results of a double-blind, placebo-controlled trial. *Hepatol Baltim Md* 12:716–724
47. Goertzel B (2010) Toward a formal characterization of real-world general intelligence. Atlantis Press
48. Groszmann RJ, Garcia-Tsao G, Bosch J, Grace ND, Burroughs AK, Planas R, Escorsell A, Garcia-Pagan JC, Patch D, Matloff DS, Gao H, Makuch R, Portal Hypertension Collaborative Group (2005) Beta-blockers to prevent gastroesophageal varices in patients with cirrhosis. *N Engl J Med* 353:2254–2261
49. Gu W, Hortlik H, Erasmus H-P, Schaaf L, Zeleke Y, Uschner FE, Ferstl P, Schulz M, Peiffer K-H, Queck A, Sauerbruch T, Brol MJ, Rohde G, Sanchez C, Moreau R, Arroyo V, Zeuzem S, Welsch C, Trebicka J (2022) Trends and the course of liver cirrhosis and its complications in Germany: Nationwide population-based study (2005 to 2018). *Lancet Reg Health – Eur* 12:
50. Guarner C, Solà R, Soriano G, Andreu M, Novella MT, Vila MC, Sàbat M, Coll S, Ortiz J, Gómez C, Balanzó J (1999) Risk of a first community-acquired spontaneous bacterial peritonitis in cirrhotics with low ascitic fluid protein levels. *Gastroenterology* 117:414–419
51. Guo A, Mazumder NR, Ladner DP, Foraker RE (2021) Predicting mortality among patients with liver cirrhosis in electronic health records with machine learning. *PLoS ONE* 16:e0256428
52. Hampe J, Cuthbert A, Croucher PJ, Mirza MM, Mascheretti S, Fisher S, Frenzel H, King K, Hasselmeyer A, MacPherson AJ, Bridger S, van Deventer S, Forbes A, Nikolaus S, Lennard-Jones JE, Foelsch UR, Krawczak M, Lewis C, Schreiber S, Mathew CG (2001) Association between insertion mutation in *NOD2* gene and Crohn's disease in German and British populations. *Lancet Lond Engl* 357:1925–1928
53. Hiraoka A, Kumada T, Michitaka K, Toyoda H, Tada T, Ueki H, Kaneto M, Aibiki T, Okudaira T, Kawakami T, Kawamura T, Yamago H, Suga Y, Miyamoto Y, Tomida H, Azemoto N, Mori K, Miyata H, Ninomiya T, Kawasaki H (2016) Usefulness of albumin–bilirubin grade for evaluation of prognosis of 2584 Japanese patients with hepatocellular carcinoma. *J Gastroenterol Hepatol* 31:1031–1036

54. Ho TK (1995) Random decision forests. IEEE Computer Society, USA
55. Hsieh W-J, Lin H-C, Hwang S-J, Hou M-C, Lee F-Y, Chang F-Y, Lee S-D (1998) The Effect of Ciprofloxacin in the Prevention of Bacterial Infection in Patients With Cirrhosis After Upper Gastrointestinal Bleeding. *Off J Am Coll Gastroenterol ACG* 93:962–966
56. Huang H, Fang M, Jostins L, Umićević Mirkov M, Boucher G, Anderson CA, Andersen V, Cleyneen I, Cortes A, Crins F, D’Amato M, Deffontaine V, Dmitrieva J, Docampo E, Elansary M, Farh KK-H, Franke A, Gori A-S, Goyette P, Halfvarson J, Haritunians T, Knight J, Lawrance IC, Lees CW, Louis E, Mariman R, Meuwissen T, Mni M, Momozawa Y, Parkes M, Spain SL, Théâtre E, Trynka G, Satsangi J, van Sommeren S, Vermeire S, Xavier RJ, International Inflammatory Bowel Disease Genetics Consortium, Weersma RK, Duerr RH, Mathew CG, Rioux JD, McGovern DPB, Cho JH, Georges M, Daly MJ, Barrett JC (2017) Fine-mapping inflammatory bowel disease loci to single-variant resolution. *Nature* 547:173–178
57. Hugot JP, Chamaillard M, Zouali H, Lesage S, Cézard JP, Belaiche J, Almer S, Tysk C, O’Morain CA, Gassull M, Binder V, Finkel Y, Cortot A, Modigliani R, Laurent-Puig P, Gower-Rousseau C, Macry J, Colombel JF, Sahbatou M, Thomas G (2001) Association of *NOD2* leucine-rich repeat variants with susceptibility to Crohn’s disease. *Nature* 411:599–603
58. Iwakiri Y, Groszmann RJ (2006) The hyperdynamic circulation of chronic liver diseases: from the patient to the molecule. *Hepatology* 43:S121-131
59. Jalan R, Fernandez J, Wiest R, Schnabl B, Moreau R, Angeli P, Stadlbauer V, Gustot T, Bernardi M, Canton R, Albillos A, Lammert F, Wilmer A, Mookerjee R, Vila J, Garcia-Martinez R, Wendon J, Such J, Cordoba J, Sanyal A, Garcia-Tsao G, Arroyo V, Burroughs A, Ginès P (2014) Bacterial infections in cirrhosis: A position statement based on the EASL Special Conference 2013. *J Hepatol* 60:1310–1324
60. Jalan R, Pavesi M, Saliba F, Amorós A, Fernandez J, Holland-Fischer P, Sawhney R, Mookerjee R, Caraceni P, Moreau R, Ginès P, Durand F, Angeli P, Alessandria C, Laleman W, Trebicka J, Samuel D, Zeuzem S, Gustot T, Gerbes AL, Wendon J, Bernardi M, Arroyo V, CANONIC Study Investigators; EASL-CLIF Consortium (2015) The CLIF Consortium Acute Decompensation score (CLIF-C ADs) for prognosis of hospitalised cirrhotic patients without acute-on-chronic liver failure. *J Hepatol* 62:831–840
61. Ji G-W, Zhu F-P, Xu Q, Wang K, Wu M-Y, Tang W-W, Li X-C, Wang X-H (2019) Machine-learning analysis of contrast-enhanced CT radiomics predicts recurrence of hepatocellular carcinoma after resection: A multi-institutional study. *EBioMedicine* 50:156–165
62. Johnson PJ, Berhane S, Kagebayashi C, Satomura S, Teng M, Reeves HL, O’Beirne J, Fox R, Skowronska A, Palmer D, Yeo W, Mo F, Lai P, Iñarrairaegui M, Chan SL, Sangro B, Miksad R, Tada T, Kumada T, Toyoda H (2015) Assessment of liver function in patients with hepatocellular carcinoma: a new evidence-based approach-the ALBI grade. *J Clin Oncol Off J Am Soc Clin Oncol* 33:550–558
63. Kamath PS, Wiesner RH, Malinchoc M, Kremers W, Therneau TM, Kosberg CL, D’Amico G, Dickson ER, Kim WR (2001) A model to predict survival in patients with end-stage liver disease. *Hepatology* 33:464–470
64. Kamath PS, Kim WR (2007) The model for end-stage liver disease (MELD). *Hepatology* 45:797–805
65. Kanwal F, Taylor TJ, Kramer JR, Cao Y, Smith D, Gifford AL, El-Serag HB, Naik AD, Asch SM (2020) Development, Validation, and Evaluation of a Simple Machine Learning Model to Predict Cirrhosis Mortality. *JAMA Netw Open* 3:e2023780

66. Kartoun U, Corey KE, Simon TG, Zheng H, Aggarwal R, Ng K, Shaw SY (2017) The MELD-Plus: A generalizable prediction risk score in cirrhosis. *PloS One* 12:e0186301
67. Kawada N, Tran-Thi TA, Klein H, Decker K (1993) The contraction of hepatic stellate (Ito) cells stimulated with vasoactive substances. Possible involvement of endothelin 1 and nitric oxide in the regulation of the sinusoidal tonus. *Eur J Biochem* 213:815–823
68. Kim MY, Baik SK, Lee SS (2010) Hemodynamic alterations in cirrhosis and portal hypertension. *Korean J Hepatol* 16:347–352
69. Kim WR, Biggins SW, Kremers WK, Wiesner RH, Kamath PS, Benson JT, Edwards E, Therneau TM (2008) Hyponatremia and mortality among patients on the liver-transplant waiting list. *N Engl J Med* 359:1018–1026
70. Kleppe A, Skrede O-J, De Raedt S, Liestøl K, Kerr DJ, Danielsen HE (2021) Designing deep learning studies in cancer diagnostics. *Nat Rev Cancer* 21:199–211
71. Konerman MA, Beste LA, Van T, Liu B, Zhang X, Zhu J, Saini SD, Su GL, Nallamotheu BK, Ioannou GN, Waljee AK (2019) Machine learning models to predict disease progression among veterans with hepatitis C virus. *PloS One* 14:e0208141
72. Krizhevsky A, Sutskever I, Hinton GE (2017) Imagenet classification with deep convolutional neural networks. *Commun ACM* 60:84–90
73. Kubesch A, Quenstedt L, Saleh M, Rüschenbaum S, Schwarzkopf K, Martinez Y, Welsch C, Zeuzem S, Welzel TM, Lange CM (2018) Vitamin D deficiency is associated with hepatic decompensation and inflammation in patients with liver cirrhosis: A prospective cohort study. *PLoS ONE* 13:e0207162
74. Lau L, Kankanige Y, Rubinstein B, Jones R, Christophi C, Muralidharan V, Bailey J (2017) Machine-Learning Algorithms Predict Graft Failure After Liver Transplantation. *Transplantation* 101:e125–e132
75. Liaw A, Wiener M (2002) Classification and regression by randomForest. *R News* 2:18–22
76. Lin Y-J, Chen R-J, Tang J-H, Yu C-S, Wu JL, Chen L-C, Chang S-S (2020) Machine-Learning Monitoring System for Predicting Mortality Among Patients With Noncancer End-Stage Liver Disease: Retrospective Study. *JMIR Med Inform* 8:e24305
77. Mitchell TM, Mitchell TM (1997) Machine learning. McGraw-hill New York
78. Montagnese S, Rautou P-E, Romero-Gómez M, Larsen FS, Shawcross DL, Thabut D, Vilstrup H, Weissenborn K (2022) EASL Clinical Practice Guidelines on the management of hepatic encephalopathy. *J Hepatol* 77:807–824
79. Moreau R, Jalan R, Gines P, Pavesi M, Angeli P, Cordoba J, Durand F, Gustot T, Saliba F, Domenicali M, Gerbes A, Wendon J, Alessandria C, Laleman W, Zeuzem S, Trebicka J, Bernardi M, Arroyo V (2013) Acute-on-Chronic Liver Failure Is a Distinct Syndrome That Develops in Patients With Acute Decompensation of Cirrhosis. *Gastroenterology* 144:1426-1437.e9
80. Murphy KP (2012) Machine learning: a probabilistic perspective. MIT press
81. Nam D, Chapiro J, Paradis V, Seraphin TP, Kather JN (2022) Artificial intelligence in liver diseases: Improving diagnostics, prognostics and response prediction. *JHEP Rep* 4:100443

82. Nischalke HD, Berger C, Aldenhoff K, Thyssen L, Gentemann M, Grünhage F, Lammert F, Nattermann J, Sauerbruch T, Spengler U, Appenrodt B (2011) Toll-like receptor (TLR) 2 promoter and intron 2 polymorphisms are associated with increased risk for spontaneous bacterial peritonitis in liver cirrhosis. *J Hepatol* 55:1010–1016
83. Novo E, Cannito S, Zamara E, Valfrè di Bonzo L, Caligiuri A, Cravanzola C, Compagnone A, Colombatto S, Marra F, Pinzani M, Parola M (2007) Proangiogenic cytokines as hypoxia-dependent factors stimulating migration of human hepatic stellate cells. *Am J Pathol* 170:1942–1953
84. Obstein KL, Campbell MS, Reddy KR, Yang Y-X (2007) Association between model for end-stage liver disease and spontaneous bacterial peritonitis. *Am J Gastroenterol* 102:2732–2736
85. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay É (2011) Scikit-learn: Machine Learning in Python. *J Mach Learn Res* 12:2825–2830
86. Peng Y, Qi X, Tang S, Deng H, Li J, Ning Z, Dai J, Hou F, Zhao J, Wang R, Guo X (2016) Child-Pugh, MELD, and ALBI scores for predicting the in-hospital mortality in cirrhotic patients with acute-on-chronic liver failure. *Expert Rev Gastroenterol Hepatol* 10:971–980
87. Pennachin C, Goertzel B (2007) Contemporary approaches to artificial general intelligence. Springer
88. Pimpin L, Cortez-Pinto H, Negro F, Corbould E, Lazarus JV, Webber L, Sheron N (2018) Burden of liver disease in Europe: Epidemiology and analysis of risk factors to identify prevention policies. *J Hepatol* 69:718–735
89. Pinato DJ, Sharma R, Allara E, Yen C, Arizumi T, Kubota K, Bettinger D, Jang JW, Smirne C, Kim YW, Kudo M, Howell J, Ramaswami R, Burlone ME, Guerra V, Thimme R, Ishizuka M, Stebbing J, Pirisi M, Carr BI (2017) The ALBI grade provides objective hepatic reserve estimation across each BCLC stage of hepatocellular carcinoma. *J Hepatol* 66:338–346
90. Powers DM (2020) Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *ArXiv Prepr ArXiv201016061*
91. Pugh RNH, Murray-Lyon IM, Dawson JL, Pietroni MC, Williams R (1973) Transection of the oesophagus for bleeding oesophageal varices. *BJS Br J Surg* 60:646–649
92. Qin H, Hu X, Zhang J, Dai H, He Y, Zhao Z, Yang J, Xu Z, Hu X, Chen Z (2021) Machine-learning radiomics to predict early recurrence in perihilar cholangiocarcinoma after curative resection. *Liver Int Off J Int Assoc Study Liver* 41:837–850
93. Quinlan JR (1986) Induction of decision trees. *Mach Learn* 1:81–106
94. Quinlan JR (2014) C4. 5: programs for machine learning. Elsevier
95. Reichert MC, Ripoll C, Casper M, Greinert R, Vandieken E, Grünhage F, Appenrodt B, Zipprich A, Lammert F (2019) Common *NOD2* Risk Variants as Major Susceptibility Factors for Bacterial Infections in Compensated Cirrhosis. *Clin Transl Gastroenterol* 10:e00002
96. Reiniš J, Petrenko O, Simbrunner B, Hofer BS, Schepis F, Scoppettuolo M, Saltini D, Indulti F, Guasconi T, Albillos A, Téllez L, Villanueva C, Brujats A, Garcia-Pagan JC, Perez-Campuzano V, Hernández-Gea V, Rautou P-E, Moga L, Vanwolleghem T, Kwanten WJ, Francque S, Trebicka J, Gu W, Ferstl PG, Gluud LL, Bendtsen F, Møller S, Kubicek S, Mandorfer M,

- Reiberger T Assessment of portal hypertension severity using machine learning models in patients with compensated cirrhosis. *J Hepatol*
97. Rezazade Mehrizi MH, van Ooijen P, Homan M (2021) Applications of artificial intelligence (AI) in diagnostic radiology: a technography study. *Eur Radiol* 31:1805–1811
 98. Ripoll C, Groszmann R, Garcia-Tsao G, Grace N, Burroughs A, Planas R, Escorsell A, Garcia-Pagan JC, Makuch R, Patch D, Matloff DS, Bosch J, Portal Hypertension Collaborative Group (2007) Hepatic venous pressure gradient predicts clinical decompensation in patients with compensated cirrhosis. *Gastroenterology* 133:481–488
 99. Rockey DC, Weisiger RA (1996) Endothelin induced contractility of stellate cells from normal and cirrhotic rat liver: implications for regulation of portal pressure and resistance. *Hepatology* 24:233–240
 100. Rolachon A, Cordier L, Bacq Y, Nousbaum JB, Franza A, Paris JC, Fratte S, Bohn B, Kitmacher P, Stahl JP (1995) Ciprofloxacin and long-term prevention of spontaneous bacterial peritonitis: results of a prospective controlled trial. *Hepatol Baltim Md* 22:1171–1174
 101. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. *Nature* 323:533–536
 102. Runyon BA (1990) Monomicrobial nonneutrocytic bacterascites: a variant of spontaneous bacterial peritonitis. *Hepatol Baltim Md* 12:710–715
 103. Russell S, Norvig P (2016) *Artificial Intelligence: A Modern Approach*. Pearson
 104. Schneider ARP, Schneider CV, Schneider KM, Baier V, Schaper S, Diedrich C, Coboeken K, Mayer H, Gu W, Trebicka J, Blank LM, Burghaus R, Lippert J, Rader DJ, Thaiss CA, Schlender J-F, Trautwein C, Kuepfer L (2022) Early prediction of decompensation (EPOD) score: Non-invasive determination of cirrhosis decompensation risk. *Liver Int* 42:640–650
 105. Schölkopf B, Smola AJ, Bach F (2002) *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press
 106. Schrier RW, Arroyo V, Bernardi M, Epstein M, Henriksen JH, Rodés J (1988) Peripheral arterial vasodilation hypothesis: a proposal for the initiation of renal sodium and water retention in cirrhosis. *Hepatology* 8:1151–1157
 107. Searle JR (1980) Minds, brains, and programs. *Behav Brain Sci* 3:417–424
 108. Spann A, Yasodhara A, Kang J, Watt K, Wang B, Goldenberg A, Bhat M (2020) Applying Machine Learning in Liver Disease and Transplantation: A Comprehensive Review. *Hepatology* 71:1093–1105
 109. Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, Montaner J, Sulkowski MS, Torriani FJ, Dieterich DT, Thomas DL, Messinger D, Nelson M (2006) Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology* 43:1317–1325
 110. Thabut D, Shah V (2010) Intrahepatic angiogenesis and sinusoidal remodeling in chronic liver disease: new targets for the treatment of portal hypertension? *J Hepatol* 53:976–980
 111. Toyoda H, Lai PB, O’Beirne J, Chong CC, Berhane S, Reeves H, Manas D, Fox RP, Yeo W, Mo F, Chan AW, Tada T, Iñarrairaegui M, Vogel A, Schweitzer N, Chan SL, Sangro B, Kumada T,

- Johnson PJ (2016) Long-term impact of liver function on curative therapy for hepatocellular carcinoma: application of the ALBI grade. *Br J Cancer* 114:744–750
112. Vilstrup H, Amodio P, Bajaj J, Cordoba J, Ferenci P, Mullen KD, Weissenborn K, Wong P (2014) Hepatic encephalopathy in chronic liver disease: 2014 Practice Guideline by the American Association for the Study Of Liver Diseases and the European Association for the Study of the Liver: Vilstrup et al. *Hepatology* 60:715–735
 113. Wang D, Wang Q, Shan F, Liu B, Lu C (2010) Identification of the risk for liver fibrosis on CHB patients using an artificial neural network based on routine and serum markers. *BMC Infect Dis* 10:251
 114. Wanless IR, Wong F, Blendis LM, Greig P, Heathcote EJ, Levy G (1995) Hepatic and portal vein thrombosis in cirrhosis: Possible role in development of parenchymal extinction and portal hypertension. *Hepatology* 21:1238–1247
 115. Wapnik WN, Tscherwonienkis AJ (1979) *Theorie der Zeichenerkennung*. De Gruyter
 116. Ward JH (1963) Hierarchical Grouping to Optimize an Objective Function. *J Am Stat Assoc* 58:236–244
 117. Wei R, Wang J, Wang X, Xie G, Wang Y, Zhang H, Peng C-Y, Rajani C, Kwee S, Liu P, Jia W (2018) Clinical prediction of HBV and HCV related hepatic fibrosis using machine learning. *EBioMedicine* 35:124–132
 118. Wiest R, Groszmann RJ (2002) The paradox of nitric oxide in cirrhosis and portal hypertension: too much, not enough. *Hepatology* 35:478–491
 119. Yasaka K, Akai H, Kunimatsu A, Abe O, Kiryu S (2018) Deep learning for staging liver fibrosis on CT: a pilot study. *Eur Radiol* 28:4578–4585
 120. Yu C-S, Chen Y-D, Chang S-S, Tang J-H, Wu JL, Lin C-H (2021) Exploring and predicting mortality among patients with end-stage liver disease without cancer: a machine learning approach. *Eur J Gastroenterol Hepatol* 33:1117–1123
 121. Zell A (1997) *Simulation neuronaler Netze*. Oldenbourg Wissenschaftsverlag
 122. Zou D, Qi X, Zhu C, Ning Z, Hou F, Zhao J, Peng Y, Li J, Deng H, Guo X (2016) Albumin-bilirubin score for predicting the in-hospital mortality of acute upper gastrointestinal bleeding in liver cirrhosis: A retrospective study. *Turk J Gastroenterol Off J Turk Soc Gastroenterol* 27:180–186

7 Publikationen

Kongressbeiträge:

38. Jahrestagung der Deutschen Arbeitsgemeinschaft zum Studium der Leber (GASL) 2022:

Müller SE, Ripoll C, Zipprich A, Bruns T, Horn P, Krawczyk M, Lammert F, Reichert MC (2022).
Machine learning models predicting decompensation in cirrhosis. Z Für Gastroenterol 60:2.17

International Liver Congress 2022:

Müller SE, Ripoll C, Zipprich A, Bruns T, Horn P, Krawczyk M, Lammert F, Reichert MC (2022).
Can machines predict liver decompensation? Analysis of 1, 415 patients with liver cirrhosis recruited
at three German referral centers. J Hepatol 77:342

→ ausgewählt zur Best Poster Presentation - Cirrhosis and Complications

35. Jahrestagung der Gastroenterologischen Arbeitsgemeinschaft Rheinland-Pfalz/Saarland (GARPS) 2022:

Müller SE, Ripoll C, Zipprich A, Bruns T, Horn P, Krawczyk M, Lammert F, Reichert MC. Können
Computer die Dekompensation einer Zirrhose vorhersagen? Eine Analyse von 1415 Patienten mit
Leberzirrhose aus drei Universitätskliniken.

→ ausgewählt für den Posterpreis Kategorie „Diagnostik“

Viszeralmedizin 2024:

Müller SE; C M; Ripoll, C; Zipprich, A; Horn, P; Krawczyk, M; Lammert, F; Reichert, MC (2024)
Vorhersage der Dekompensation einer Leberzirrhose mithilfe maschinellen Lernens. Z Für
Gastroenterol 62:KV 123

Zeitschriftenbeiträge:

Müller SE, Casper M, Ripoll C, Zipprich A, Horn P, Krawczyk M, Lammert F, Reichert MC. Machine learning models predicting decompensation in cirrhosis.

→ eingereicht bei: Journal of Gastrointestinal and Liver Diseases

Dank

An erster Stelle möchte ich meinem Doktorvater und Betreuer, Herrn Prof. Dr. med. Dipl. Volkswirt Frank Lammert, meinen aufrichtigen Dank aussprechen. Durch seine fachliche Expertise, die Überlassung des spannenden Themas und seine Unterstützung während der gesamten Promotionszeit hat er maßgeblich zum Gelingen dieser Arbeit beigetragen.

Ebenso gebührt Herrn PD Dr. Matthias Reichert mein tiefster Dank, der mich kontinuierlich bei allen Fragen, insbesondere bei der Arbeit mit der INCA-Datenbank, unterstützt hat. Seine schnellen und hilfreichen Antworten auf meine Fragen sowie seine wertvollen Anmerkungen waren für den Fortschritt dieser Arbeit von unschätzbarem Wert.

Mein aufrichtiger Dank gilt auch dem gesamten Team der INCA-Studie in Homburg, Halle und Jena, das mir den Zugang zu den gesammelten Daten ermöglicht hat. Ohne ihre umfangreiche Datenbasis wäre diese Arbeit nicht möglich gewesen.

Mein größter Dank jedoch gilt meiner Familie.

Lebenslauf

Aus datenschutzrechtlichen Gründen wird der Lebenslauf in der elektronischen Fassung der Dissertation nicht veröffentlicht.

Anhang:

1. Entscheidungsbaum

- a. Retrospektive Beurteilung
 - i. Ungeclustert

Tabelle 26: Entscheidungsbaum ungeclustert retrospektiv

Albumin	5,065
Bilirubin Max	3,160
rs5743271 = -1	1,873
AFP Max	1,730
rs72796367 = -1	1,519
TAG Max	0,668
Bilirubin	0,643
CRP	0,519
PTT	0,264
Ätiologie = 1	0,251
Na Min	0,138
Harnstoff	0,105
CHE Min	0,091
Albumin Min	0,089
Ammoniak Max	0,088
AST	0,081
rs5743271 = 0	0,053
ALT	0,042
ALT Max	0,032
Leukozyten	0,029
Leukozyten Min	0,027
Thrombozyten	0,027
Harnstoff Max	0,021
AP Max	0,021
INR Max	0,018
Varizen = 1	0,017
Varizen = 0	0,016
Kreatinin	0,015
Langfristige Einnahme ATB = 0	0,012
Hb	0,010
Ca Max	0,009
Glucose Max	0,008
Kreatinin Max	0,007
Ferritin Min	0,005
Thrombozyten Min	0,004

K Max	0,004
Ferritin Max	0,003
CRP Max	0,002
Diabetes Mellitus = 1	0,001
Cholesterol Max	0,001
Betablocker = 1	0,000
Geschlecht = 1	0,000
ACLF Baseline = -1	0,000
ACLF Baseline = 0	0,000
ACLF Baseline = 1	0,000
ACLF Baseline = 2	0,000
ACLF Baseline = 3	0,000
rs5743271 = 1	0,000
rs72796367 = 0	0,000
rs72796367 = 1	0,000
rs72796367 = 2	0,000
p.R702W = -1	0,000
p.R702W = 0	0,000
p.R702W = 1	0,000
p.R702W = 2	0,000
p.G908R = -1	0,000
p.G908R = 0	0,000
p.G908R = 1	0,000
p.G908R = 2	0,000
c.3020insC = -1	0,000
c.3020insC = 0	0,000
c.3020insC = 1	0,000
c.3020insC = 2	0,000
Diabetes Mellitus = 0	0,000
Ätiologie = 2	0,000
Ätiologie = 3	0,000
Ätiologie = 5	0,000
Ätiologie = 6	0,000
Ätiologie = 7	0,000
Ätiologie = 8	0,000
Ätiologie = 9	0,000
Ätiologie = 10	0,000
Ätiologie = 11	0,000
Varizen = -1	0,000
PPI = -1	0,000
PPI = 0	0,000
Betablocker = -1	0,000
Betablocker = 0	0,000
Langfristige Einnahme ATB = -1	0,000

Langfristige Einnahme ATB = 1	0,000
PTT Max	0,000
Ca	0,000
Geschlecht = 0	0,000
INR	-0,002
Ca Min	-0,002
PPI = 1	-0,004
HB Min	-0,005
GGT Max	-0,005
K Min	-0,005
Na	-0,007
AST Max	-0,008
Ätiologie = 4	-0,008
K	-0,011

ii. Geclustert

Tabelle 27: Entscheidungsbaum geclustert retrospektiv

Albumin	4,763
Bilirubin Max	3,435
AFP Max	2,679
rs5743271 = -1	1,586
rs72796367 = -1	1,073
Bilirubin	0,767
CRP	0,538
Albumin Min	0,308
TAG Max	0,290
CRP Max	0,210
Na Min	0,162
Ätiologie = 1	0,133
ALT	0,099
INR	0,081
Ammoniak Max	0,079
Varizen = 0	0,070
Hb	0,044
AP Max	0,041
Varizen = 1	0,035
Ferritin Min	0,032
Leukozyten Min	0,028
Langfristige Einnahme ATB = 1	0,021
AST	0,019
Kreatinin	0,017
CHE Min	0,015
K Max	0,012
Leukozyten	0,008
Betablocker = 0	0,008
Thrombozyten	0,008
ALT Max	0,007
INR Max	0,004
Ätiologie = 4	0,003
Ca Max	0,003
K Min	0,002
Kreatinin Max	0,002
Ferritin Max	0,001
PPI = 1	0,001
Glucose Max	0,001
p.R702W = 1	0,001

PPI = 0	0,000
Geschlecht = 0	0,000
ACLF Baseline = -1	0,000
ACLF Baseline = 0	0,000
ACLF Baseline = 1	0,000
ACLF Baseline = 2	0,000
ACLF Baseline = 3	0,000
rs5743271 = 0	0,000
rs5743271 = 1	0,000
rs72796367 = 0	0,000
rs72796367 = 1	0,000
rs72796367 = 2	0,000
p.R702W = -1	0,000
p.R702W = 0	0,000
p.R702W = 2	0,000
p.G908R = -1	0,000
p.G908R = 0	0,000
p.G908R = 1	0,000
p.G908R = 2	0,000
c.3020insC = -1	0,000
c.3020insC = 0	0,000
c.3020insC = 1	0,000
c.3020insC = 2	0,000
Diabetes Mellitus = 0	0,000
Diabetes Mellitus = 1	0,000
Ätiologie = 2	0,000
Ätiologie = 3	0,000
Ätiologie = 5	0,000
Ätiologie = 6	0,000
Ätiologie = 7	0,000
Ätiologie = 8	0,000
Ätiologie = 9	0,000
Ätiologie = 10	0,000
Ätiologie = 11	0,000
Varizen = -1	0,000
PPI = -1	0,000
Betablocker = -1	0,000
Langfristige Einnahme ATB = -1	0,000
Langfristige Einnahme ATB = 0	0,000
Thrombozyten Min	-0,001
Cholesterol Max	-0,006
GGT Max	-0,006
Betablocker = 1	-0,006
Geschlecht = 1	-0,006

Na	-0,007
K	-0,007

b. Prospektive Beurteilung

i. Ungeclustert

Tabelle 28: Entscheidungsbaum ungeclustert prospektiv

Bilirubin Max	4,408
Harnstoff Max	2,055
Albumin	0,777
CHE Min	0,701
Kreatinin Max	0,373
Ammoniak Max	0,230
INR	0,184
Bilirubin	0,124
CRP	0,110
CRP Max	0,090
PTT	0,081
Albumin Min	0,067
ALT Max	0,047
Harnstoff	0,046
Ferritin Min	0,025
Varizen = 0	0,018
HB Min	0,018
Ca Min	0,017
Na Min	0,016
Ätiologie = 1	0,012
Ca Max	0,012
AP Max	0,012
ALT	0,007
TAG Max	0,006
Geschlecht = 0	0,000
Geschlecht = 1	0,000
ACLF Baseline = -1	0,000
ACLF Baseline = 0	0,000
ACLF Baseline = 1	0,000
ACLF Baseline = 2	0,000
ACLF Baseline = 3	0,000
rs5743271 = -1	0,000
rs5743271 = 0	0,000
rs5743271 = 1	0,000
rs72796367 = -1	0,000
rs72796367 = 0	0,000
rs72796367 = 1	0,000
rs72796367 = 2	0,000
p.R702W = -1	0,000

p.R702W = 0	0,000
p.R702W = 1	0,000
p.G908R = -1	0,000
p.G908R = 0	0,000
p.G908R = 1	0,000
c.3020insC = -1	0,000
c.3020insC = 0	0,000
c.3020insC = 1	0,000
c.3020insC = 2	0,000
Diabetes Mellitus = 0	0,000
Diabetes Mellitus = 1	0,000
Ätiologie = 2	0,000
Ätiologie = 3	0,000
Ätiologie = 4	0,000
Ätiologie = 5	0,000
Ätiologie = 6	0,000
Ätiologie = 7	0,000
Ätiologie = 8	0,000
Ätiologie = 9	0,000
Ätiologie = 10	0,000
Ätiologie = 11	0,000
Varizen = -1	0,000
PPI = -1	0,000
PPI = 0	0,000
PPI = 1	0,000
Betablocker = -1	0,000
Betablocker = 0	0,000
Betablocker = 1	0,000
Langfristige Einnahme ATB = -1	0,000
Langfristige Einnahme ATB = 0	0,000
Langfristige Einnahme ATB = 1	0,000
K Max	0,000
Kreatinin	0,000
AFP Max	0,000
Leukozyten Min	0,000
Thrombozyten Min	0,000
Ca	-0,004
GGT Max	-0,006
Varizen = 1	-0,007
PTT Max	-0,007
Ferritin Max	-0,007
Cholesterol Max	-0,008
K	-0,008
AST	-0,008

Na	-0,008
AST Max	-0,011
INR Max	-0,012
Thrombozyten	-0,013
Hb	-0,015
K Min	-0,021
Leukozyten	-0,024
Glucose Max	-0,045

ii. Geclustert

Tabelle 29: Entscheidungsbaum geclustert prospektiv

Bilirubin Max	4,326
Kreatinin Max	2,652
INR	0,867
CHE Min	0,768
Albumin	0,693
Albumin Min	0,324
Ammoniak Max	0,263
Varizen = 1	0,136
CRP	0,118
Kreatinin	0,114
Bilirubin	0,098
CRP Max	0,073
Glucose Max	0,042
Varizen = 0	0,041
AFP Max	0,041
Ferritin Min	0,030
TAG Max	0,028
Thrombozyten	0,018
Leukozyten Min	0,017
Thrombozyten Min	0,017
Hb	0,015
Cholesterol Max	0,012
Leukozyten	0,012
AP Max	0,007
INR Max	0,002
K Min	0,002
Ferritin Max	0,000
Geschlecht = 0	0,000
Geschlecht = 1	0,000
ACLF Baseline = -1	0,000
ACLF Baseline = 0	0,000
ACLF Baseline = 1	0,000
ACLF Baseline = 2	0,000
ACLF Baseline = 3	0,000
rs5743271 = -1	0,000
rs5743271 = 0	0,000
rs5743271 = 1	0,000
rs72796367 = -1	0,000
rs72796367 = 0	0,000

rs72796367 = 1	0,000
rs72796367 = 2	0,000
p.R702W = -1	0,000
p.R702W = 0	0,000
p.R702W = 1	0,000
p.G908R = -1	0,000
p.G908R = 0	0,000
p.G908R = 1	0,000
c.3020insC = -1	0,000
c.3020insC = 0	0,000
c.3020insC = 1	0,000
c.3020insC = 2	0,000
Diabetes Mellitus = 0	0,000
Diabetes Mellitus = 1	0,000
Ätiologie = 1	0,000
Ätiologie = 2	0,000
Ätiologie = 3	0,000
Ätiologie = 4	0,000
Ätiologie = 5	0,000
Ätiologie = 6	0,000
Ätiologie = 7	0,000
Ätiologie = 8	0,000
Ätiologie = 9	0,000
Ätiologie = 10	0,000
Ätiologie = 11	0,000
Varizen = -1	0,000
PPI = -1	0,000
PPI = 0	0,000
PPI = 1	0,000
Betablocker = -1	0,000
Betablocker = 0	0,000
Betablocker = 1	0,000
Langfristige Einnahme ATB = -1	0,000
Langfristige Einnahme ATB = 0	0,000
Langfristige Einnahme ATB = 1	0,000
Ca Max	0,000
ALT Max	-0,001
ALT	-0,008
K Max	-0,010
K	-0,011
Na Min	-0,029
GGT Max	-0,035
AST	-0,035
Na	-0,042

2. Random Forest

a. Retrospektive Beurteilung

i. Ungeclustert

Tabelle 30: Random Forest ungeclustert retrospektiv

Bilirubin Max	0,784
Albumin	0,778
Bilirubin	0,346
CRP	0,280
Varizen = 0	0,113
AFP Max	0,111
Hb	0,103
Varizen = 1	0,096
rs5743271 = -1	0,090
ALT Max	0,086
rs72796367 = -1	0,069
Ca	0,062
Na Min	0,058
Harnstoff	0,053
INR	0,049
Ätiologie = 1	0,049
Ferritin Min	0,045
AP Max	0,044
AST	0,042
GGT Max	0,039
ALT	0,038
rs72796367 = 0	0,036
Langfristige Einnahme ATB = 1	0,034
rs5743271 = 0	0,030
Na	0,028
Kreatinin	0,027
PTT	0,026
Cholesterol Max	0,025
Albumin Min	0,025
Ferritin Max	0,023
Ammoniak Max	0,023
ACLF Baseline = 0	0,022
Langfristige Einnahme ATB = 0	0,021
AST Max	0,017
Leukozyten Min	0,015
Thrombozyten	0,015
TAG Max	0,014

CRP Max	0,013
Leukozyten	0,013
Ätiologie = 4	0,008
Geschlecht = 1	0,006
Ätiologie = 2	0,005
Diabetes Mellitus = 1	0,004
ACLF Baseline = 1	0,003
p.R702W = 1	0,002
Betablocker = 1	0,002
Thrombozyten Min	0,002
PPI = 1	0,001
Geschlecht = 0	0,001
c.3020insC = 0	0,001
Ätiologie = 3	0,001
Ätiologie = 10	0,001
p.R702W = 0	0,000
K Max	0,000
p.G908R = 1	0,000
c.3020insC = 1	0,000
Langfristige Einnahme ATB = -1	0,000
ACLF Baseline = -1	0,000
ACLF Baseline = 3	0,000
rs72796367 = 2	0,000
p.R702W = -1	0,000
p.R702W = 2	0,000
p.G908R = -1	0,000
p.G908R = 2	0,000
c.3020insC = -1	0,000
c.3020insC = 2	0,000
Ätiologie = 6	0,000
Ätiologie = 8	0,000
Ätiologie = 9	0,000
Ätiologie = 7	0,000
Varizen = -1	0,000
p.G908R = 0	0,000
PPI = -1	0,000
Betablocker = -1	0,000
rs5743271 = 1	0,000
ACLF Baseline = 2	0,000
rs72796367 = 1	-0,001
Ätiologie = 11	-0,001
Ätiologie = 5	-0,001
Diabetes Mellitus = 0	-0,001
Ca Max	-0,001

PPI = 0	-0,002
Betablocker = 0	-0,003
CHE Min	-0,005
Glucose Max	-0,005
K	-0,006
INR Max	-0,007
K Min	-0,010
PTT Max	-0,012
Harnstoff Max	-0,022
Kreatinin Max	-0,023
Ca Min	-0,025
HB Min	-0,026

ii. Geclustert

Tabelle 31: Random Forest geclustert retrospektiv

Albumin	0,757
Bilirubin Max	0,677
Bilirubin	0,389
CRP	0,346
AFP Max	0,133
Hb	0,132
Varizen = 0	0,105
ALT Max	0,090
rs72796367 = -1	0,081
Varizen = 1	0,075
rs5743271 = -1	0,071
INR	0,064
ALT	0,063
Ätiologie = 1	0,057
Kreatinin	0,051
Na Min	0,045
Langfristige Einnahme ATB = 0	0,042
Langfristige Einnahme ATB = 1	0,040
rs72796367 = 0	0,037
Ferritin Min	0,035
GGT Max	0,034
Albumin Min	0,033
Ferritin Max	0,033
rs5743271 = 0	0,032
Na	0,029
Cholesterol Max	0,023
AST	0,022
Ammoniak Max	0,020
ACLF Baseline = 0	0,020
CHE Min	0,018
CRP Max	0,017
K	0,009
Leukozyten Min	0,007
Diabetes Mellitus = 1	0,006
Ätiologie = 4	0,006
Geschlecht = 1	0,006
ACLF Baseline = 1	0,005
Ätiologie = 2	0,005
Leukozyten	0,005

TAG Max	0,005
Thrombozyten	0,005
Glucose Max	0,004
Ätiologie = 10	0,004
Diabetes Mellitus = 0	0,004
PPI = 1	0,003
p.R702W = 0	0,003
Geschlecht = 0	0,003
K Max	0,003
Ca Max	0,003
K Min	0,002
p.R702W = 1	0,002
Betablocker = 1	0,002
Betablocker = 0	0,001
c.3020insC = 0	0,001
c.3020insC = 1	0,000
Ätiologie = 7	0,000
Ätiologie = 11	0,000
ACLF Baseline = 2	0,000
Varizen = -1	0,000
PPI = -1	0,000
Langfristige Einnahme ATB = -1	0,000
ACLF Baseline = -1	0,000
ACLF Baseline = 3	0,000
rs72796367 = 2	0,000
p.R702W = -1	0,000
p.R702W = 2	0,000
p.G908R = -1	0,000
p.G908R = 2	0,000
c.3020insC = -1	0,000
c.3020insC = 2	0,000
Ätiologie = 6	0,000
Ätiologie = 8	0,000
Ätiologie = 9	0,000
Ätiologie = 3	0,000
rs5743271 = 1	0,000
Betablocker = -1	0,000
Ätiologie = 5	0,000
p.G908R = 0	0,000
p.G908R = 1	0,000
AP Max	-0,001
PPI = 0	-0,001
rs72796367 = 1	-0,001
INR Max	-0,002

Kreatinin Max	-0,008
Thrombozyten Min	-0,016

b. Prospektive Beurteilung

i. Ungeclustert

Tabelle 32: Random Forest ungeclustert prospektiv

Bilirubin Max	0,577
Albumin	0,463
Kreatinin Max	0,370
Bilirubin	0,262
CRP	0,230
Harnstoff Max	0,224
PTT	0,223
Ammoniak Max	0,190
PTT Max	0,164
Albumin Min	0,126
ALT Max	0,096
INR Max	0,091
Hb	0,088
INR	0,085
CHE Min	0,085
Varizen = 1	0,082
TAG Max	0,082
K Min	0,079
Na	0,074
Ca Min	0,073
AP Max	0,070
Ätiologie = 1	0,069
AST	0,059
HB Min	0,057
Varizen = 0	0,047
Thrombozyten	0,047
Leukozyten	0,040
Langfristige Einnahme ATB = 1	0,029
Ca	0,024
Diabetes Mellitus = 0	0,023
Thrombozyten Min	0,023
Geschlecht = 1	0,022
K	0,022
Harnstoff	0,021
Betablocker = 0	0,015
Langfristige Einnahme ATB = 0	0,015
Kreatinin	0,010
p.R702W = 1	0,006

Diabetes Mellitus = 1	0,006
Betablocker = 1	0,006
Ätiologie = 2	0,005
Ätiologie = 10	0,005
CRP Max	0,004
Glucose Max	0,002
Ferritin Min	0,002
p.R702W = 0	0,001
Ätiologie = 4	0,001
PPI = 1	0,000
ACLF Baseline = -1	0,000
ACLF Baseline = 2	0,000
ACLF Baseline = 3	0,000
rs5743271 = 1	0,000
rs72796367 = 1	0,000
rs72796367 = 2	0,000
p.R702W = -1	0,000
p.G908R = -1	0,000
p.G908R = 0	0,000
p.G908R = 1	0,000
c.3020insC = -1	0,000
c.3020insC = 0	0,000
c.3020insC = 1	0,000
c.3020insC = 2	0,000
Ätiologie = 3	0,000
Ätiologie = 5	0,000
Ätiologie = 6	0,000
Ätiologie = 7	0,000
Ätiologie = 8	0,000
Ätiologie = 9	0,000
Ätiologie = 11	0,000
Varizen = -1	0,000
PPI = -1	0,000
Betablocker = -1	0,000
Langfristige Einnahme ATB = -1	0,000
PPI = 0	-0,001
rs5743271 = 0	-0,002
ACLF Baseline = 1	-0,006
Ca Max	-0,006
ACLF Baseline = 0	-0,007
K Max	-0,011
GGT Max	-0,013
ALT	-0,013
Na Min	-0,013

rs72796367 = 0	-0,015
Leukozyten Min	-0,016
Geschlecht = 0	-0,018
Ferritin Max	-0,025
Cholesterol Max	-0,028
rs5743271 = -1	-0,030
rs72796367 = -1	-0,047
AST Max	-0,095
AFP Max	-0,105

ii. Geclustert

Tabelle 33: Random Forest geclustert prospektiv

Bilirubin Max	0,899
Kreatinin Max	0,423
Albumin	0,301
CRP	0,262
Bilirubin	0,246
Albumin Min	0,223
Ammoniak Max	0,156
INR	0,139
Varizen = 1	0,124
INR Max	0,107
CHE Min	0,107
Na	0,103
Varizen = 0	0,098
AP Max	0,093
K Min	0,087
Ferritin Min	0,081
Ätiologie = 1	0,050
Kreatinin	0,045
Glucose Max	0,036
Hb	0,035
Langfristige Einnahme ATB = 1	0,032
AST	0,032
Ferritin Max	0,028
Leukozyten	0,025
ALT	0,022
Leukozyten Min	0,017
Langfristige Einnahme ATB = 0	0,016
Geschlecht = 1	0,015
ACLF Baseline = 1	0,015
Diabetes Mellitus = 0	0,015
ACLF Baseline = 0	0,013
Thrombozyten Min	0,013
K	0,012
PPI = 0	0,005
PPI = 1	0,004
Ca Max	0,004
p.G908R = 0	0,001
Ätiologie = 11	0,001
Betablocker = 0	0,001

Na Min	0,001
ACLF Baseline = -1	0,000
ACLF Baseline = 2	0,000
ACLF Baseline = 3	0,000
rs5743271 = 1	0,000
rs72796367 = 1	0,000
rs72796367 = 2	0,000
p.R702W = -1	0,000
p.R702W = 0	0,000
p.G908R = -1	0,000
p.G908R = 1	0,000
c.3020insC = -1	0,000
c.3020insC = 0	0,000
c.3020insC = 1	0,000
c.3020insC = 2	0,000
Ätiologie = 3	0,000
Ätiologie = 5	0,000
Ätiologie = 6	0,000
Ätiologie = 7	0,000
Ätiologie = 8	0,000
Ätiologie = 9	0,000
Varizen = -1	0,000
PPI = -1	0,000
Betablocker = -1	0,000
Langfristige Einnahme ATB = -1	0,000
Ätiologie = 4	-0,001
Thrombozyten	-0,001
Ätiologie = 10	-0,002
Geschlecht = 0	-0,004
p.R702W = 1	-0,006
Ätiologie = 2	-0,006
Diabetes Mellitus = 1	-0,007
K Max	-0,007
Cholesterol Max	-0,011
rs72796367 = 0	-0,012
ALT Max	-0,012
GGT Max	-0,012
Betablocker = 1	-0,019
rs72796367 = -1	-0,022
CRP Max	-0,028
rs5743271 = -1	-0,030
TAG Max	-0,032
rs5743271 = 0	-0,041
AFP Max	-0,084

3. SVM

a. Retrospektive Beurteilung

i. Ungeclustert

Tabelle 34: SVM ungeclustert retrospektiv

Bilirubin	0,997
Albumin	0,821
Na Min	0,572
Albumin Min	0,286
AST	0,284
rs72796367 = -1	0,271
rs5743271 = -1	0,271
Ferritin Min	0,270
rs5743271 = 0	0,249
Ätiologie = 1	0,217
Kreatinin	0,217
rs72796367 = 0	0,216
Leukozyten Min	0,190
CRP Max	0,175
K Max	0,173
Varizen = 1	0,169
Varizen = 0	0,169
Ca	0,167
HB Min	0,154
Bilirubin Max	0,152
ALT Max	0,146
Langfristige Einnahme ATB = 1	0,143
K Min	0,131
Ammoniak Max	0,126
Hb	0,099
Langfristige Einnahme ATB = 0	0,090
GGT Max	0,084
AP Max	0,080
CHE Min	0,079
PTT	0,079
Ferritin Max	0,075
Na	0,070
Ätiologie = 2	0,068
Leukozyten	0,067
Ca Min	0,065
ALT	0,060
Ätiologie = 10	0,046

Ca Max	0,042
INR Max	0,039
p.R702W = 1	0,032
Thrombozyten	0,027
INR	0,026
Ätiologie = 4	0,025
c.3020insC = 1	0,020
p.R702W = 0	0,019
TAG Max	0,016
ACLF Baseline = 1	0,012
PTT Max	0,011
Harnstoff	0,011
CRP	0,010
Ätiologie = 3	0,008
c.3020insC = 0	0,007
ACLF Baseline = 0	0,007
Thrombozyten Min	0,006
AST Max	0,003
ACLF Baseline = 2	0,001
Diabetes Mellitus = 0	0,001
Ätiologie = 11	0,001
Ätiologie = 9	0,001
Diabetes Mellitus = 1	0,001
rs72796367 = 2	0,000
c.3020insC = 2	0,000
ACLF Baseline = -1	0,000
p.R702W = 2	0,000
ACLF Baseline = 3	0,000
p.G908R = 2	0,000
p.R702W = -1	-0,001
p.G908R = -1	-0,001
c.3020insC = -1	-0,001
Cholesterol Max	-0,001
Ätiologie = 6	-0,001
Geschlecht = 1	-0,002
Geschlecht = 0	-0,002
Varizen = -1	-0,003
PPI = 1	-0,004
PPI = -1	-0,004
Betablocker = -1	-0,004
Ätiologie = 8	-0,004
p.G908R = 0	-0,005
Langfristige Einnahme ATB = -1	-0,005
p.G908R = 1	-0,005

rs5743271 = 1	-0,006
Ätiologie = 5	-0,008
Ätiologie = 7	-0,009
rs72796367 = 1	-0,014
PPI = 0	-0,014
K	-0,017
Harnstoff Max	-0,019
Betablocker = 1	-0,025
Betablocker = 0	-0,028
Kreatinin Max	-0,031
Glucose Max	-0,032
AFP Max	-0,034

ii. Geclustert

Tabelle 35: SVM geclustert retrospektiv

Bilirubin	1,117
Albumin	0,815
Na Min	0,586
AST	0,381
rs5743271 = -1	0,330
rs72796367 = -1	0,324
rs5743271 = 0	0,318
Ferritin Min	0,296
ALT Max	0,272
Kreatinin	0,271
rs72796367 = 0	0,242
CRP Max	0,227
Varizen = 1	0,212
Ätiologie = 1	0,200
Albumin Min	0,195
Varizen = 0	0,187
Hb	0,157
Bilirubin Max	0,145
Leukozyten Min	0,143
Ammoniak Max	0,122
K Max	0,120
Langfristige Einnahme ATB = 1	0,116
GGT Max	0,110
CHE Min	0,087
Langfristige Einnahme ATB = 0	0,079
Ätiologie = 10	0,077
Ätiologie = 2	0,077
Ca Max	0,072
INR	0,067
ALT	0,066
Na	0,064
Leukozyten	0,060
K Min	0,047
Ferritin Max	0,046
INR Max	0,046
AP Max	0,045
p.R702W = 1	0,025
c.3020insC = 1	0,022
K	0,018
Cholesterol Max	0,016
Ätiologie = 4	0,013

Ätiologie = 3	0,013
ACLF Baseline = 1	0,011
TAG Max	0,008
Glucose Max	0,008
p.R702W = 0	0,007
Diabetes Mellitus = 0	0,007
Diabetes Mellitus = 1	0,007
Ätiologie = 11	0,007
Thrombozyten	0,005
c.3020insC = 0	0,004
Ätiologie = 5	0,003
Geschlecht = 0	0,002
Geschlecht = 1	0,002
ACLF Baseline = 2	0,002
p.G908R = 1	0,001
Ätiologie = 9	0,001
ACLF Baseline = 0	0,001
p.G908R = 0	0,000
ACLF Baseline = 3	0,000
ACLF Baseline = -1	0,000
c.3020insC = 2	0,000
rs72796367 = 2	0,000
p.G908R = 2	0,000
p.R702W = 2	0,000
p.R702W = -1	-0,001
p.G908R = -1	-0,001
c.3020insC = -1	-0,001
PPI = 1	-0,001
Ätiologie = 8	-0,001
Ätiologie = 6	-0,002
Langfristige Einnahme ATB = -1	-0,003
Varizen = -1	-0,003
PPI = -1	-0,003
Betablocker = -1	-0,003
rs72796367 = 1	-0,005
rs5743271 = 1	-0,008
Ätiologie = 7	-0,008
PPI = 0	-0,009
Betablocker = 0	-0,018
Betablocker = 1	-0,018
AFP Max	-0,018
CRP	-0,018
Kreatinin Max	-0,025
Thrombozyten Min	-0,039

- b. Prospektive Beurteilung
i. Ungeclustert

Tabelle 36: SVM ungeclustert prospektiv

Bilirubin Max	0,772
Na	0,657
Albumin	0,481
PTT	0,364
Ammoniak Max	0,353
CRP	0,325
rs5743271 = -1	0,292
rs72796367 = -1	0,288
rs72796367 = 0	0,284
Harnstoff Max	0,273
CHE Min	0,252
rs5743271 = 0	0,244
Kreatinin Max	0,239
Albumin Min	0,205
PTT Max	0,172
K Min	0,164
AP Max	0,150
Varizen = 1	0,097
Ätiologie = 1	0,090
INR Max	0,080
Thrombozyten Min	0,079
AST Max	0,074
Varizen = 0	0,068
Hb	0,068
TAG Max	0,059
Geschlecht = 0	0,055
GGT Max	0,055
Ferritin Min	0,052
Geschlecht = 1	0,051
K Max	0,048
p.R702W = 0	0,046
ALT Max	0,044
PPI = 1	0,040
Bilirubin	0,038
Kreatinin	0,038
p.R702W = 1	0,030
Betablocker = 1	0,028
Leukozyten Min	0,028

PPI = 0	0,011
Ätiologie = 2	0,008
Betablocker = 0	0,008
CRP Max	0,007
Na Min	0,006
HB Min	0,005
Ätiologie = 5	0,004
Langfristige Einnahme ATB = 0	0,002
rs5743271 = 1	0,001
Ätiologie = 8	0,001
Ätiologie = 10	0,001
Ca Max	0,001
Ätiologie = 11	0,000
ACLF Baseline = -1	0,000
ACLF Baseline = 2	0,000
ACLF Baseline = 3	0,000
rs72796367 = 2	0,000
p.R702W = -1	0,000
p.G908R = -1	0,000
p.G908R = 0	0,000
p.G908R = 1	0,000
c.3020insC = -1	0,000
c.3020insC = 2	0,000
Ätiologie = 9	0,000
Ätiologie = 3	0,000
ACLF Baseline = 1	-0,001
rs72796367 = 1	-0,001
Ätiologie = 7	-0,001
ALT	-0,001
Cholesterol Max	-0,002
Thrombozyten	-0,002
ACLF Baseline = 0	-0,005
Varizen = -1	-0,005
c.3020insC = 1	-0,006
Ätiologie = 6	-0,006
Langfristige Einnahme ATB = 1	-0,006
Ferritin Max	-0,007
Ätiologie = 4	-0,008
AST	-0,008
PPI = -1	-0,012
Betablocker = -1	-0,012
Langfristige Einnahme ATB = -1	-0,012
c.3020insC = 0	-0,013
Ca Min	-0,019

Diabetes Mellitus = 0	-0,022
INR	-0,028
K	-0,028
Diabetes Mellitus = 1	-0,029
AFP Max	-0,033
Leukozyten	-0,034
Ca	-0,039
Glucose Max	-0,047
Harnstoff	-0,087

ii. Geclustert

Tabelle 37: SVM geclustert prospektiv

Bilirubin Max	0,865
Na	0,759
Albumin	0,742
CRP	0,326
Ammoniak Max	0,247
INR Max	0,224
Kreatinin Max	0,215
CHE Min	0,170
Ätiologie = 1	0,139
INR	0,135
rs5743271 = -1	0,126
rs72796367 = -1	0,124
rs72796367 = 0	0,099
rs5743271 = 0	0,097
Hb	0,072
Albumin Min	0,063
TAG Max	0,039
PPI = 1	0,023
Geschlecht = 1	0,010
Ferritin Min	0,010
Ätiologie = 3	0,008
c.3020insC = 1	0,007
PPI = -1	0,005
Betablocker = -1	0,005
Langfristige Einnahme ATB = -1	0,005
p.G908R = 1	0,004
Bilirubin	0,004
Ätiologie = 10	0,002
ACLF Baseline = 1	0,001
ACLF Baseline = 2	0,001
rs5743271 = 1	0,001
Ätiologie = 8	0,001
ACLF Baseline = -1	0,000
ACLF Baseline = 3	0,000
rs72796367 = 2	0,000
p.R702W = -1	0,000
p.G908R = -1	0,000
c.3020insC = -1	0,000
c.3020insC = 2	0,000

Ätiologie = 9	0,000
Ätiologie = 7	0,000
AP Max	0,000
rs72796367 = 1	-0,001
Betablocker = 0	-0,001
p.G908R = 0	-0,002
Ätiologie = 6	-0,002
Varizen = -1	-0,002
c.3020insC = 0	-0,005
PPI = 0	-0,005
Geschlecht = 0	-0,006
Ätiologie = 4	-0,007
Ätiologie = 5	-0,008
Diabetes Mellitus = 0	-0,011
Thrombozyten Min	-0,013
p.R702W = 0	-0,016
AST	-0,016
ACLF Baseline = 0	-0,017
Ätiologie = 11	-0,017
p.R702W = 1	-0,023
Ätiologie = 2	-0,023
Leukozyten	-0,023
Langfristige Einnahme ATB = 0	-0,024
Langfristige Einnahme ATB = 1	-0,024
Diabetes Mellitus = 1	-0,025
Betablocker = 1	-0,030
Thrombozyten	-0,033
Glucose Max	-0,039
ALT	-0,051
K Min	-0,052
Ferritin Max	-0,053
K	-0,069
Cholesterol Max	-0,076
Kreatinin	-0,079
ALT Max	-0,082
Na Min	-0,091
Varizen = 0	-0,095
Varizen = 1	-0,097
AFP Max	-0,102
K Max	-0,113
GGT Max	-0,114
Ca Max	-0,125
Leukozyten Min	-0,141
CRP Max	-0,155

4. Neuronales Netz

a. Retrospektive Beurteilung

i. Ungeclustert

Tabelle 38: Neuronales Netz ungeclustert retrospektiv

Bilirubin	1,596
Albumin	0,682
Na Min	0,540
Albumin Min	0,530
rs72796367 = -1	0,456
Kreatinin	0,382
Ferritin Min	0,373
rs5743271 = -1	0,364
Bilirubin Max	0,358
Langfristige Einnahme ATB = 1	0,304
ALT Max	0,296
Varizen = 1	0,268
AST	0,258
rs72796367 = 0	0,252
Ca	0,249
AP Max	0,248
rs5743271 = 0	0,246
Leukozyten Min	0,226
Varizen = 0	0,215
Hb	0,185
K Min	0,184
Ätiologie = 1	0,179
HB Min	0,152
PTT	0,152
CRP Max	0,150
Ammoniak Max	0,136
CHE Min	0,136
Leukozyten	0,106
Ca Max	0,106
K Max	0,090
ALT	0,089
INR Max	0,086
GGT Max	0,077
Ferritin Max	0,075
Langfristige Einnahme ATB = 0	0,073
c.3020insC = 1	0,073
Ca Min	0,072

CRP	0,064
Thrombozyten Min	0,061
INR	0,061
ACLF Baseline = 1	0,059
Harnstoff	0,053
Ätiologie = 3	0,050
Kreatinin Max	0,046
p.R702W = 1	0,043
Geschlecht = 1	0,042
Cholesterol Max	0,039
Na	0,035
PTT Max	0,035
TAG Max	0,031
PPI = 0	0,028
Ätiologie = 2	0,025
Geschlecht = 0	0,023
AST Max	0,023
Harnstoff Max	0,023
Diabetes Mellitus = 0	0,022
Betablocker = 1	0,022
PPI = 1	0,020
Betablocker = 0	0,020
ACLF Baseline = 0	0,019
ACLF Baseline = 2	0,018
Ätiologie = 4	0,015
Thrombozyten	0,014
p.R702W = 0	0,014
Ätiologie = 9	0,014
Ätiologie = 10	0,012
Ätiologie = 11	0,010
p.G908R = 1	0,010
Langfristige Einnahme ATB = -1	0,006
Varizen = -1	0,005
Diabetes Mellitus = 1	0,002
Ätiologie = 8	0,001
rs72796367 = 2	0,000
p.R702W = 2	0,000
p.G908R = 2	0,000
c.3020insC = 2	0,000
ACLF Baseline = -1	0,000
c.3020insC = 0	-0,001
p.R702W = -1	-0,001
ACLF Baseline = 3	-0,003
PPI = -1	-0,004

c.3020insC = -1	-0,004
Betablocker = -1	-0,004
p.G908R = -1	-0,005
Glucose Max	-0,008
Ätiologie = 6	-0,010
Ätiologie = 5	-0,010
p.G908R = 0	-0,012
Ätiologie = 7	-0,018
rs72796367 = 1	-0,026
rs5743271 = 1	-0,027
AFP Max	-0,030
K	-0,072

ii. Geclustert

Tabelle 39: Neuronales Netz geclustert retrospektiv

Bilirubin	1,594
Albumin	0,822
Na Min	0,618
rs5743271 = -1	0,460
rs72796367 = -1	0,451
Bilirubin Max	0,394
Albumin Min	0,380
Ferritin Min	0,370
ALT Max	0,365
Kreatinin	0,363
rs5743271 = 0	0,348
Varizen = 1	0,311
AST	0,302
Langfristige Einnahme ATB = 1	0,295
rs72796367 = 0	0,256
Hb	0,229
Varizen = 0	0,228
AP Max	0,208
K Max	0,199
K Min	0,196
Ätiologie = 1	0,195
CHE Min	0,161
GGT Max	0,151
Leukozyten	0,148
CRP Max	0,147
Ca Max	0,133
Ferritin Max	0,112
INR	0,107
Ammoniak Max	0,105
Leukozyten Min	0,104
Ätiologie = 3	0,082
TAG Max	0,075
Thrombozyten Min	0,073
INR Max	0,073
Langfristige Einnahme ATB = 0	0,071
c.3020insC = 1	0,064
Cholesterol Max	0,057
ALT	0,055
Glucose Max	0,052

Ätiologie = 2	0,051
CRP	0,050
Thrombozyten	0,039
Diabetes Mellitus = 1	0,037
p.R702W = 1	0,032
Ätiologie = 10	0,030
ACLF Baseline = 1	0,029
Diabetes Mellitus = 0	0,028
Ätiologie = 11	0,018
Geschlecht = 1	0,017
Betablocker = 0	0,015
PPI = 1	0,013
Ätiologie = 9	0,013
Geschlecht = 0	0,013
Ätiologie = 8	0,012
Ätiologie = 4	0,011
ACLF Baseline = 3	0,001
p.G908R = 1	0,001
Na	0,001
p.G908R = 2	0,000
c.3020insC = 2	0,000
ACLF Baseline = -1	0,000
rs72796367 = 2	0,000
p.R702W = 2	0,000
ACLF Baseline = 0	0,000
p.R702W = 0	-0,002
p.R702W = -1	-0,005
ACLF Baseline = 2	-0,005
Betablocker = -1	-0,005
Varizen = -1	-0,006
c.3020insC = -1	-0,008
p.G908R = -1	-0,009
Betablocker = 1	-0,009
p.G908R = 0	-0,010
PPI = -1	-0,010
Ätiologie = 6	-0,011
Ätiologie = 5	-0,011
Langfristige Einnahme ATB = -1	-0,012
c.3020insC = 0	-0,014
Ätiologie = 7	-0,018
rs72796367 = 1	-0,019
rs5743271 = 1	-0,021
Kreatinin Max	-0,022
PPI = 0	-0,033

K	-0,044
AFP Max	-0,058

b. Prospektive Beurteilung

i. Ungeclustert

Tabelle 40: Neuronales Netz ungeclustert prospektiv

Bilirubin Max	0,670
Na	0,481
PTT	0,343
Albumin	0,286
CRP	0,227
rs5743271 = -1	0,205
rs72796367 = -1	0,189
CHE Min	0,184
rs72796367 = 0	0,156
Bilirubin	0,153
K Min	0,141
rs5743271 = 0	0,122
Geschlecht = 1	0,114
Kreatinin Max	0,105
c.3020insC = 1	0,103
INR Max	0,082
AST Max	0,078
Ammoniak Max	0,072
ACLF Baseline = 1	0,067
TAG Max	0,064
Ätiologie = 1	0,052
Albumin Min	0,018
Varizen = 1	0,012
ALT Max	0,010
Geschlecht = 0	0,006
ACLF Baseline = 0	0,005
ACLF Baseline = 3	0,002
ACLF Baseline = 2	0,001
Ätiologie = 3	0,001
Ätiologie = 9	0,001
Ätiologie = 6	0,000
ACLF Baseline = -1	0,000
rs72796367 = 2	0,000
c.3020insC = 2	0,000
p.R702W = -1	-0,001
p.G908R = -1	-0,001
c.3020insC = -1	-0,001
Ätiologie = 8	-0,002

rs5743271 = 1	-0,004
p.G908R = 0	-0,004
c.3020insC = 0	-0,008
Ätiologie = 4	-0,013
rs72796367 = 1	-0,017
PTT Max	-0,017
Ätiologie = 10	-0,019
Langfristige Einnahme ATB = 1	-0,021
Ätiologie = 5	-0,023
PPI = -1	-0,024
Ferritin Min	-0,027
ALT	-0,030
p.G908R = 1	-0,032
Thrombozyten Min	-0,035
Varizen = -1	-0,036
Harnstoff Max	-0,036
Betablocker = -1	-0,038
Langfristige Einnahme ATB = -1	-0,038
Diabetes Mellitus = 1	-0,040
p.R702W = 0	-0,042
Langfristige Einnahme ATB = 0	-0,045
AP Max	-0,045
Ätiologie = 11	-0,046
Ca Min	-0,048
Ätiologie = 2	-0,051
HB Min	-0,051
Diabetes Mellitus = 0	-0,052
AFP Max	-0,056
Leukozyten	-0,056
Ätiologie = 7	-0,062
Betablocker = 1	-0,064
INR	-0,068
Varizen = 0	-0,069
Leukozyten Min	-0,070
Glucose Max	-0,078
Ca	-0,078
AST	-0,078
Betablocker = 0	-0,080
PPI = 1	-0,081
Ferritin Max	-0,082
Ca Max	-0,085
PPI = 0	-0,088
Kreatinin	-0,092
p.R702W = 1	-0,098

Na Min	-0,099
Harnstoff	-0,101
CRP Max	-0,104
K	-0,108
Cholesterol Max	-0,114
Hb	-0,118
GGT Max	-0,130
Thrombozyten	-0,147
K Max	-0,158

ii. Geclustert

Tabelle 41: Neuronales Netz geclustert prospektiv

Bilirubin Max	0,885
Na	0,641
Albumin	0,413
rs72796367 = -1	0,376
rs5743271 = -1	0,366
CRP	0,314
K Min	0,291
rs72796367 = 0	0,268
TAG Max	0,252
Ätiologie = 1	0,246
rs5743271 = 0	0,225
INR Max	0,213
Ammoniak Max	0,210
Kreatinin Max	0,152
Bilirubin	0,113
CHE Min	0,088
c.3020insC = 1	0,085
AP Max	0,064
Geschlecht = 1	0,061
PPI = 0	0,059
Albumin Min	0,056
Ätiologie = 10	0,052
Ferritin Min	0,048
INR	0,036
Varizen = 1	0,034
ACLF Baseline = 1	0,029
ACLF Baseline = 2	0,024
Ätiologie = 3	0,019
Thrombozyten	0,012
Ätiologie = 8	0,010
Ätiologie = 5	0,007
rs5743271 = 1	0,004
Ätiologie = 6	0,004
p.R702W = -1	0,002
p.G908R = -1	0,002
c.3020insC = 0	0,002
Betablocker = 0	0,002
ALT Max	0,002
c.3020insC = -1	0,001
ACLF Baseline = -1	0,000

rs72796367 = 2	0,000
c.3020insC = 2	0,000
Ätiologie = 9	0,000
ACLF Baseline = 3	0,000
ACLF Baseline = 0	-0,001
Ferritin Max	-0,002
Langfristige Einnahme ATB = 1	-0,004
Ätiologie = 11	-0,005
PPI = 1	-0,010
Hb	-0,010
Ätiologie = 2	-0,012
Langfristige Einnahme ATB = 0	-0,012
Langfristige Einnahme ATB = -1	-0,013
PPI = -1	-0,016
Betablocker = -1	-0,016
K Max	-0,017
p.G908R = 0	-0,019
Varizen = -1	-0,022
Ätiologie = 7	-0,024
Diabetes Mellitus = 1	-0,027
rs72796367 = 1	-0,028
Betablocker = 1	-0,028
Ätiologie = 4	-0,032
Ca Max	-0,032
ALT	-0,036
Cholesterol Max	-0,038
AST	-0,039
Geschlecht = 0	-0,041
p.R702W = 0	-0,044
Leukozyten	-0,050
Diabetes Mellitus = 0	-0,056
Varizen = 0	-0,056
p.G908R = 1	-0,064
Na Min	-0,064
AFP Max	-0,072
p.R702W = 1	-0,081
Kreatinin	-0,087
Glucose Max	-0,107
Leukozyten Min	-0,108
CRP Max	-0,115
K	-0,132
GGT Max	-0,153
Thrombozyten Min	-0,161