

qx-Club | 01.10.2019

DEEP LEARNING TECHNIKEN IM EINSATZ

Anwendungen im Kontext von Economic Capital und Cash Flow
Projektionsmodellen

Götz Cypra | UNIQA Insurance Group

Dr. Mario Hörig, Dr. Daniel Hohmann | Oliver Wyman Actuarial Services



Inhalt

1.	Hintergrund und Motivation	3
2.	Zur Theorie Neuronaler Netze In a Nutshell	8
3.	Erste Anwendungsbeispiele	17
4.	Optimierung der SAA unter Solvency II Risikokennzahlen Case Study	22
5.	Erkenntnisse und Ausblick	36

Hintergrund und Motivation

Hintergrund und Motivation

Anwendungshinweise

- Deep Learning Verfahren, wie neuronale Netze (NN) sind – sofern richtig eingesetzt mächtige Werkzeuge mit vielversprechenden Eigenschaften
- Dennoch sind sie keine Kristallkugeln oder allwissende Wundermaschinen, welche mit minimalem Aufwand trainiert werden und dann später für alle möglichen Szenarien eine belastbare Aussage treffen können.
Auch für NN wirken die „Schwerkkräfte der Statistik“ weiterhin
- Man muss in der Lage sein, hinreichend viele und gute Daten über den Zusammenhang zwischen Zielgrößen und erklärenden Variablen zu sammeln oder erzeugen
- Die Frage ist, wo der Sweet Spot von NN liegt und wir sie dort einsetzen können



Hintergrund und Motivation

Abstrakte Aufgabenstellung

- Zielgröße Y soll erklärt werden durch erklärende Variablen x_1, \dots, x_n via funktionalem Zusammenhang (x_i nennen wir von nun auch „Risikofaktor“)
 - Beispiele für Kombinationen aus Zielgröße und erklärenden Variablen:
 - Solvency II - Eigenmittel zu $t = 0$ erklärt durch Marktbedingungen zu $t = 0$ und/oder Assetallokations-Parameter
 - Eigenmittel zu $t = j, j > 0$, Marktbedingungen zu $t = 1, \dots, j$
 - Cash Flows eines Versicherungsbestands zu $t = j$ und dafür relevante erklärende Variablen
-

Vorteile von Deep Learning Verfahren gegenüber herkömmlichen Verfahren

1. Bessere Abbildung von gemischten Effekten
2. Einfachere und kanonischere Abbildungen von Nebenbedingungen oder nicht-glaten Effekten (z.B. falls Zielvariable nicht negativ werden darf oder einen „Knick“ aufweist)
3. Dynamische Form des Lernens möglich

Alle drei Aspekte werden später anhand separater Beispiele illustriert.

Vorgehen

Trainingsszenarien

- „Viele“ Initiale Stresse der „Risikofaktoren“, zwischen denen und deren Einfluss auf die Eigenmittel ein Zusammenhang hergestellt werden soll



- **Konkret:**

- Marktrisikofaktoren (Gestalt der initialen Zinskurve, Spreadlandschaften, Marktwerte von Aktien/Immobilien)
- Parameter der initialen Assetallokation

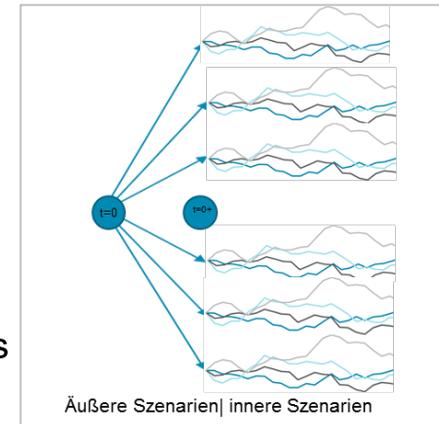
- Risikoneutrale Fortschreibung dieser initialen Marktbedingungen pro äußerem Stress

- Dabei Verwendung von nur „wenigen“ Pfaden zur Bewertung, über die dann die Eigenmittel pro äußerem Stress gemittelt werden.



- **Idee:**

- Sammele an möglichst vielen Stellen im hochdimensionalen Risikofaktorraum Information über den Zusammenhang zwischen Risikofaktoren und Eigenmitteln
- Informationen sind ungenau (hohe Varianz) aber unbiased.
- NN kann idealerweise den Trend daraus extrahieren ohne den Noise mitzuerklären.

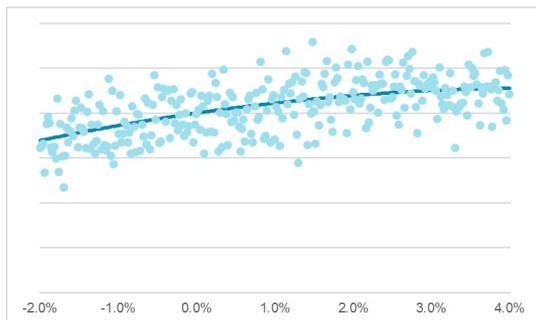


Hintergrund und Motivation

Gründe für die Wahl neuronaler Netze

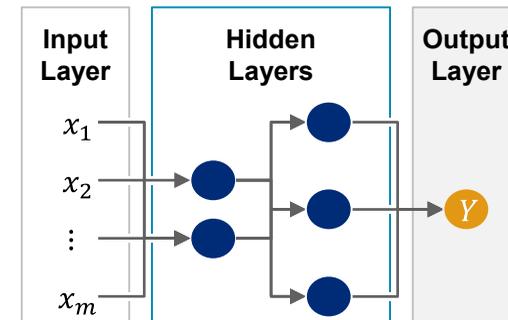
Vorgehen in UNIQA derzeit

- Derzeit Least Square Monte Carlo (LSMC) bei UNIQA für SCR-Berechnung im Einsatz
- Konkret: Polynomiale Erklärung des Zusammenhangs zwischen Risikofaktoren und Eigenmittel unter Verwendung von Modellwahlverfahren
- Fokus liegt auf SCR-Berechnung - Miteinbeziehung der Änderung der Assetallokation nicht im Fokus



Erwünschte Verbesserungen mit Neuronalen Netzen

- Bessere Abbildung durch höhere Sensitivität gegenüber Änderungen Assetallokation
- Komplexere Zusammenhänge abbildbar
- Flexibler: neue und andere Inputs durch Training integrierbar, insb. Asset übergreifende Kennzahlen
- Kontinuierliches Trainieren auf freien Serverkapazitäten und mit regulären Läufen
- Längere Nutzbarkeit durch komplexere Zusammenhänge und Nachtraining



Zur Theorie Neuronaler Netze

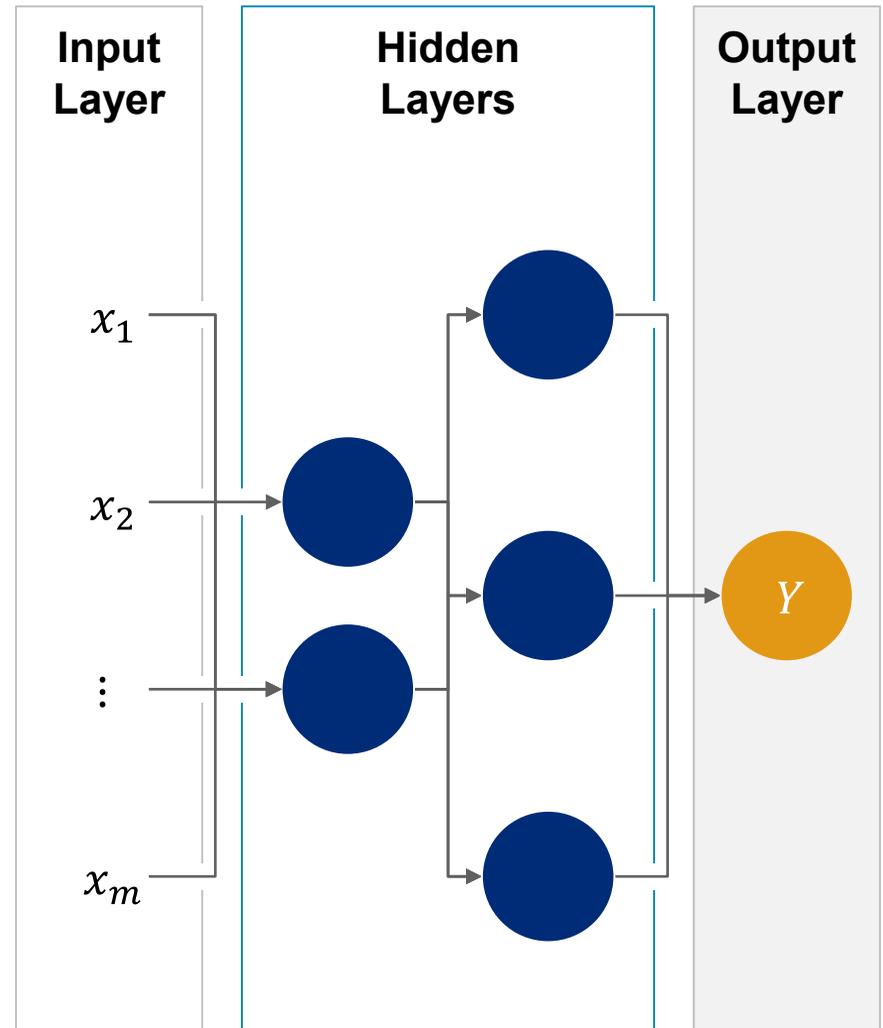
In a Nutshell

Zur Theorie Neuronaler Netze

Überblick

- Ein neuronales Netz definiert eine Funktion zur Signalverarbeitung, die sich in Form eines azyklischen, gerichteten Graphen darstellen lässt
- Eingangs- und Ausgangssignal können beliebige Dimension besitzen
- Die eigentlichen Rechenoperationen finden innerhalb der „Neuronen“ statt
- Klassische Anwendungsgebiete von neuronalen Netzen sind Bild-, Text- und Spracherkennung
- Wir werden neuronale Netze ausschließlich im Kontext der klassischen Regressionen verwenden:

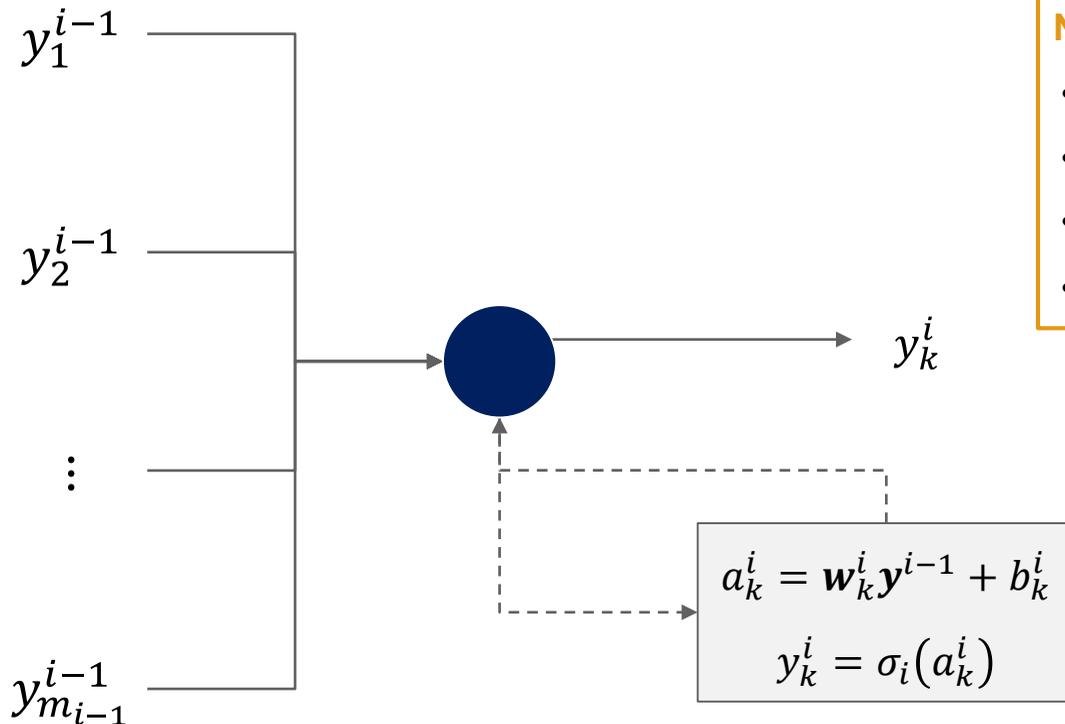
$$Y = NN(x_1, \dots, x_m) + \varepsilon$$



Zur Theorie Neuronaler Netze

Das Neuron

Rechenoperationen innerhalb eines Neurons



Notation

- y_k^i : Ausgabe von Neuron k in Layer i
- $y_k^0 := x_k$
- m_i : Anzahl Neuronen in Layer i
- \mathbf{w}_k^i, b_k^i : Parameter von Neuron k in Layer i

Die elementaren Rechenoperationen innerhalb jedes einzelnen Neurons zusammen mit einer bekannten Netzstruktur machen das NN erstaunlich handhabbar

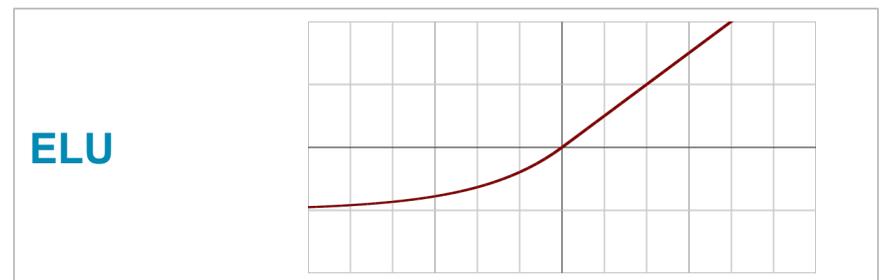
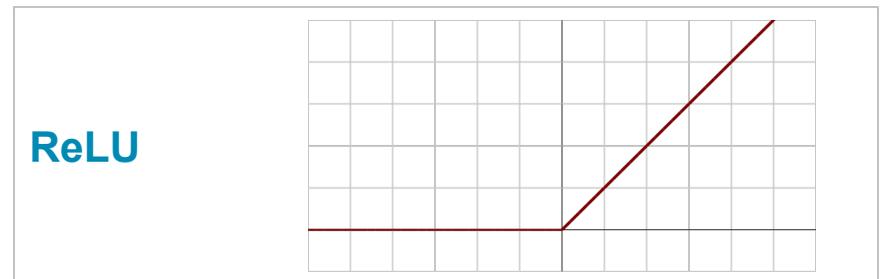
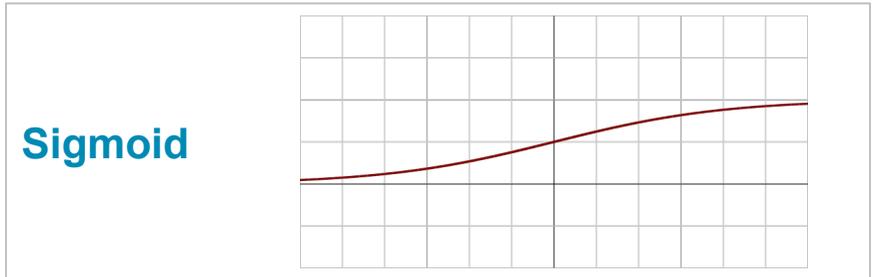
Zur Theorie Neuronaler Netze

Aktivierungsfunktionen

- Aktivierungsfunktionen sind notwendig, um Nichtlinearitäten abbilden zu können:

$$\mathbf{y}^i = \sigma_i(\mathbf{W}^i \sigma_{i-1}(\mathbf{W}^{i-1} \mathbf{y}^{i-2} + \mathbf{b}^{i-1}) + \mathbf{b}^i)$$

- Die Aktivierungsfunktionen können sich von Layer zu Layer unterscheiden
- Die Aktivierungsfunktion des Ausgabelayers hat wesentlichen Einfluss auf die Form und das Verhalten des Ausgabesignals
- Zur Funktionsapproximation verwenden wir im Ausgabelayer i.d.R. die Identität (lineare Aktivierungsfunktion)



Die Aktivierungsfunktionen sind der wesentliche Bestandteil für die Erzeugung komplexer Abhängigkeitsstrukturen

Zur Theorie Neuronaler Netze

Warum „Deep Learning“

Die Tiefe des neuronalen Netzes hat erheblichen Einfluss auf das Verhältnis zwischen Kapazität und Größe des Hypothesenraums

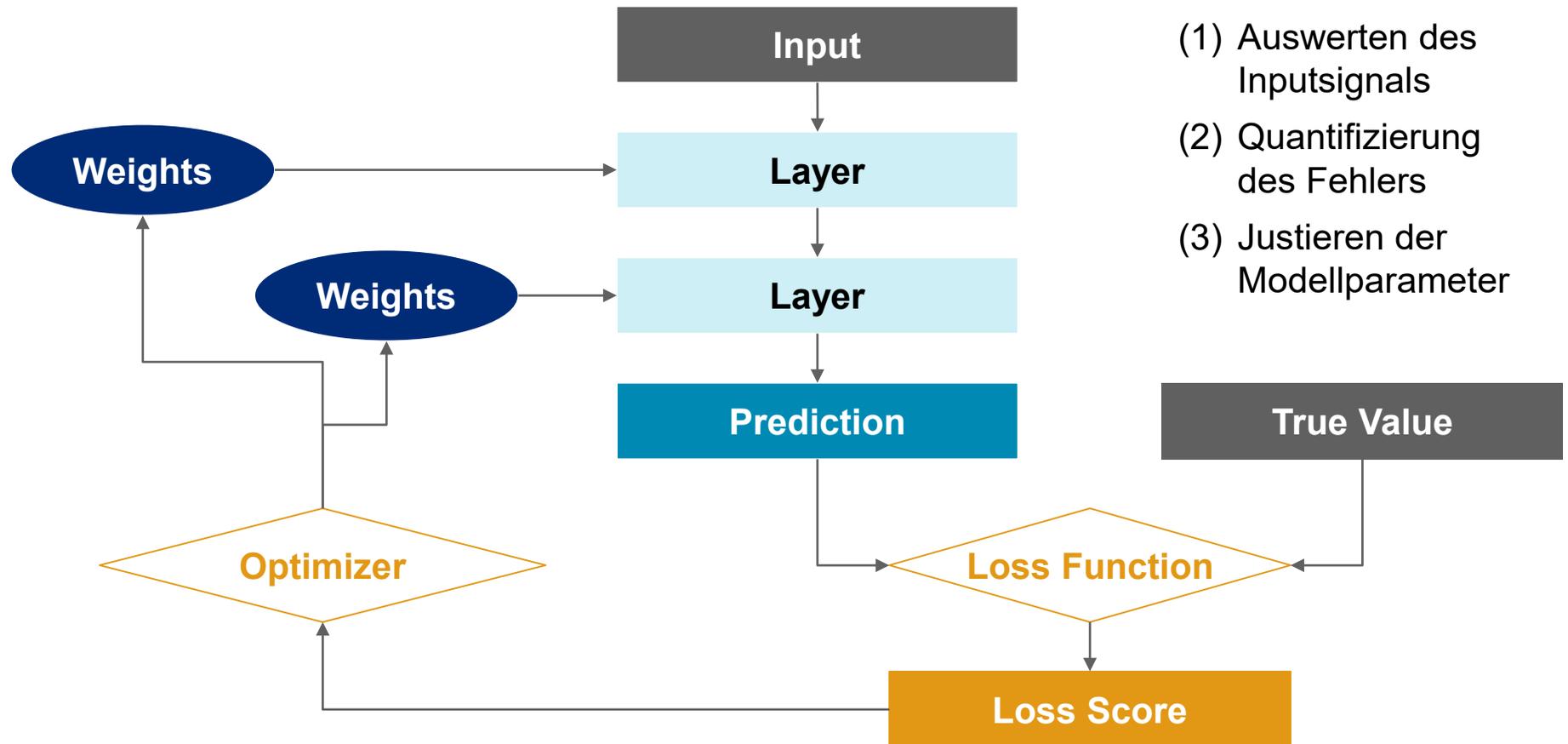
Shallow neural networks require exponentially more neurons than deep networks to achieve comparable function approximation precisions

(Liang and Skrikant, 2017)

Deep Learning ermöglicht damit eine effizientere Abbildung komplexer Abhängigkeitsstrukturen

Zur Theorie Neuronaler Netze

Der Trainingszyklus



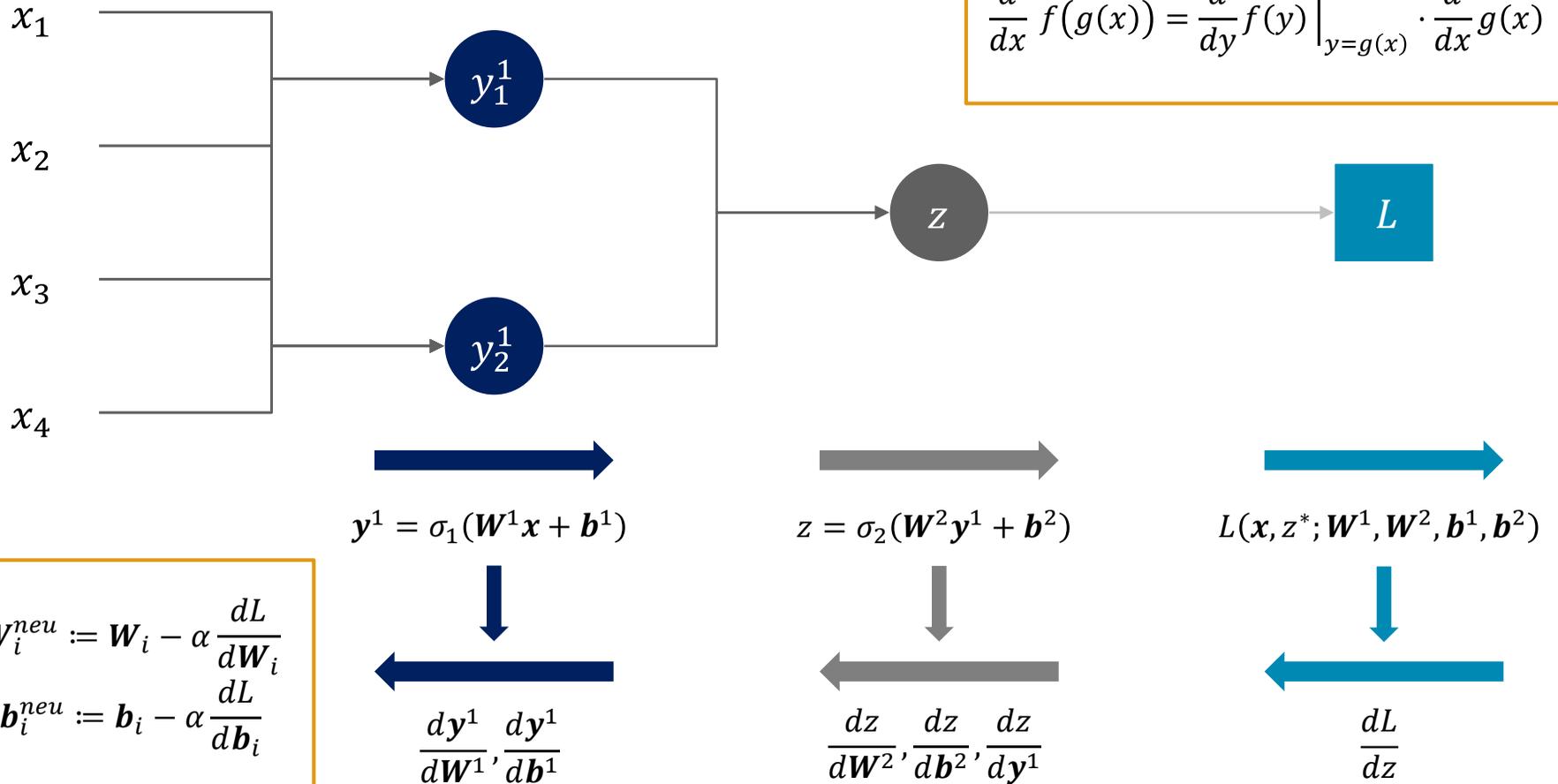
Für die Kalibrierung der Modellparameter ist ein vielfaches Durchlaufen des Trainingszyklus notwendig

Zur Theorie Neuronaler Netze

Optimierung der Parameter mittels Backpropagation

Kettenregel der Ableitung:

$$\frac{d}{dx} f(g(x)) = \frac{d}{dy} f(y) \Big|_{y=g(x)} \cdot \frac{d}{dx} g(x)$$



Zur Theorie Neuronaler Netze

Hyperparameter

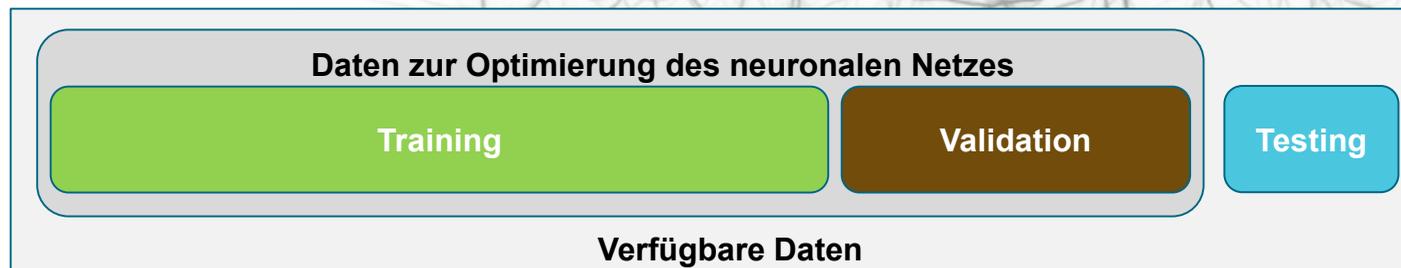
Eine Vielzahl von Hyperparametern können nicht adaptiv in das Trainieren des neuronalen Netzes mit einbezogen werden:

Struktur des NN

- Anzahl Layer
- Anzahl Neuronen pro Layer
- Aktivierungsfunktion pro Layer
- Abhängigkeitsstruktur

Lernprozess

- Verlustfunktion
- Optimierer
- Schrittweite
- Batch Size
- Anzahl Epochen



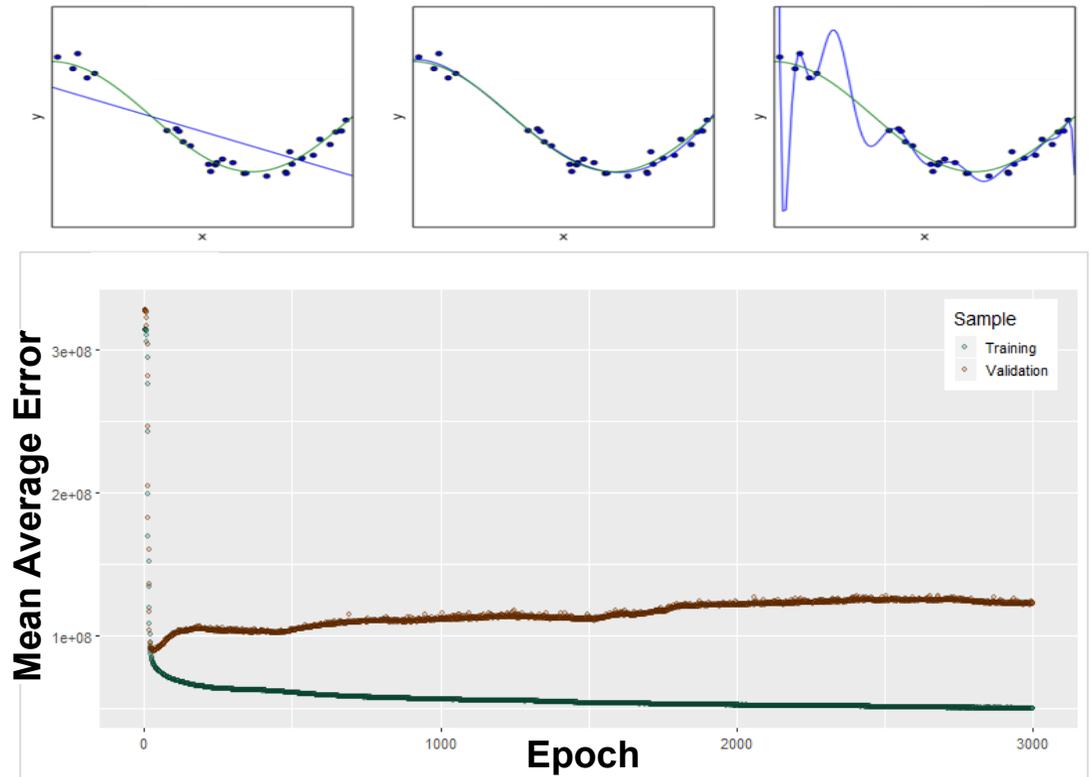
Die „optimale“ Struktur des NN wird durch Auswerten des trainierten NN auf dem Validation Sample ermittelt.

Zur Theorie Neuronaler Netze

Overfitting und Regularisierung

Neuronale Netze sind anfällig für Overfitting. Verbreitete Ansätze zur Regularisierung bei Deep Learning:

- Weight Regularization
- Dropout
- Early Stopping
- Reduktion der Komplexität (Layer, Anzahl Neuronen) des Netzes



Durch ihre Eigenschaft, komplexe funktionale Zusammenhänge abbilden zu können, sind Neuronale Netze anfällig für Overfitting.

Erste Anwendungsbeispiele

Erste Anwendungsbeispiele

Erklärung des Barwerts der zukünftigen Überschüsse (FDB)

Aufgabenstellung

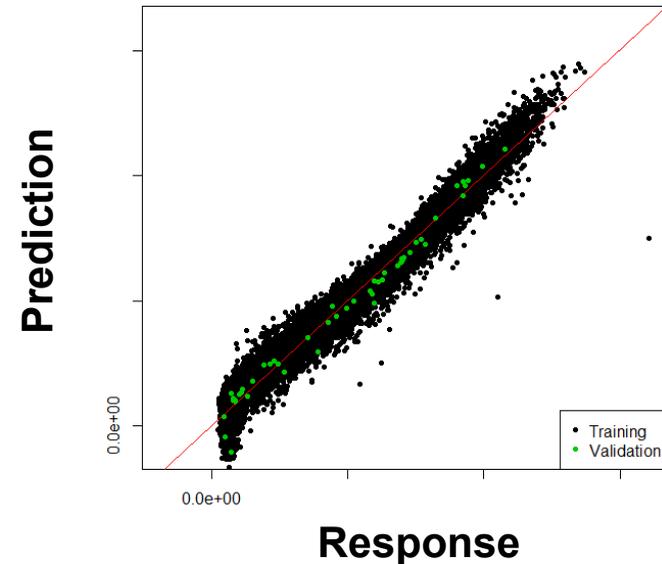
Erkläre den Wert der FDB unter Änderung der kompletten Marktrisiko-Landschaft (parametrisiert durch 17 Risikofaktoren)

Herausforderungen

- Finden eines 17-dimensionalen Zusammenhangs zwischen den Marktrisikofaktoren und einer Zielgröße generell hochdimensional
- FDB ist Größe mit hochkomplexem Zusammenhang zum Kapitalmarkt
- FDB kann nicht negativ werden

Polynomialer LSMC Ansatz

Automatische Modellwahl, ca. 200 Einzelterme inkl. gemischten Termen



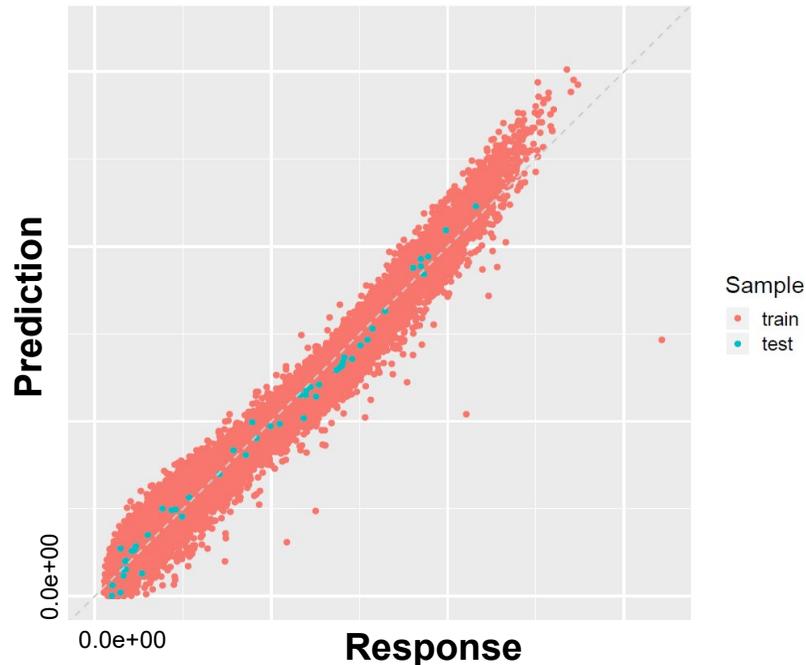
Scatterplot bei polynomialen Ansatz weist „lokalen Bias“ in Wellenform auf, welcher sich aus Fehlen von Termen und Problemen mit unterer Grenze bei Wert von Null ergibt

Erste Anwendungsbeispiele

Erklärung des Barwerts der zukünftigen Überschüsse (FDB)

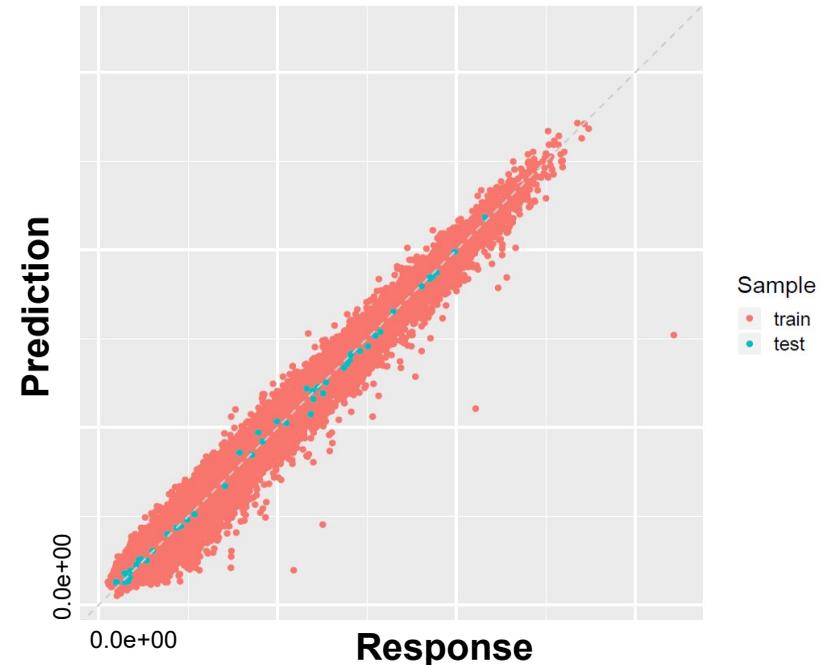
Flaches Neuronales Netz

- 1 Layer à 3'000 Neuronen (Kapazität 57'001)
- 2'000 Epochen Training



Tiefes Neuronales Netz

- 5 Layer à 5 Neuronen (Kapazität 216)
- 200 Epochen Training



Tiefes NN liefert einen wesentlich besseren Fit (trotz deutlich geringerer Komplexität und weniger Trainingsdurchläufen), während NN mit 1 Layer bereits mit Wertuntergrenze umgehen kann

Erste Anwendungsbeispiele

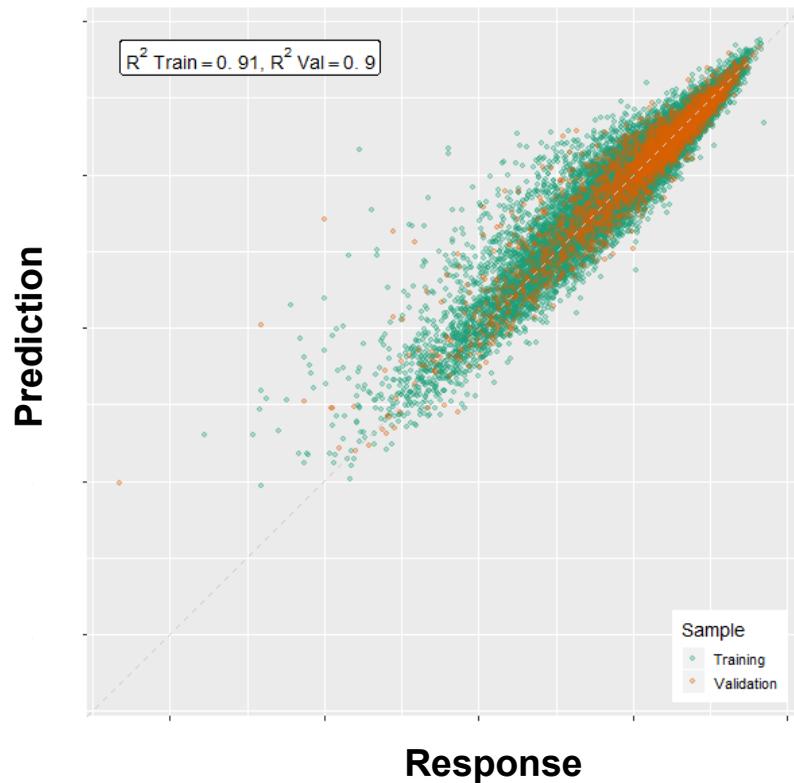
Erklärung des PVFP durch graduelles Training

Run 1

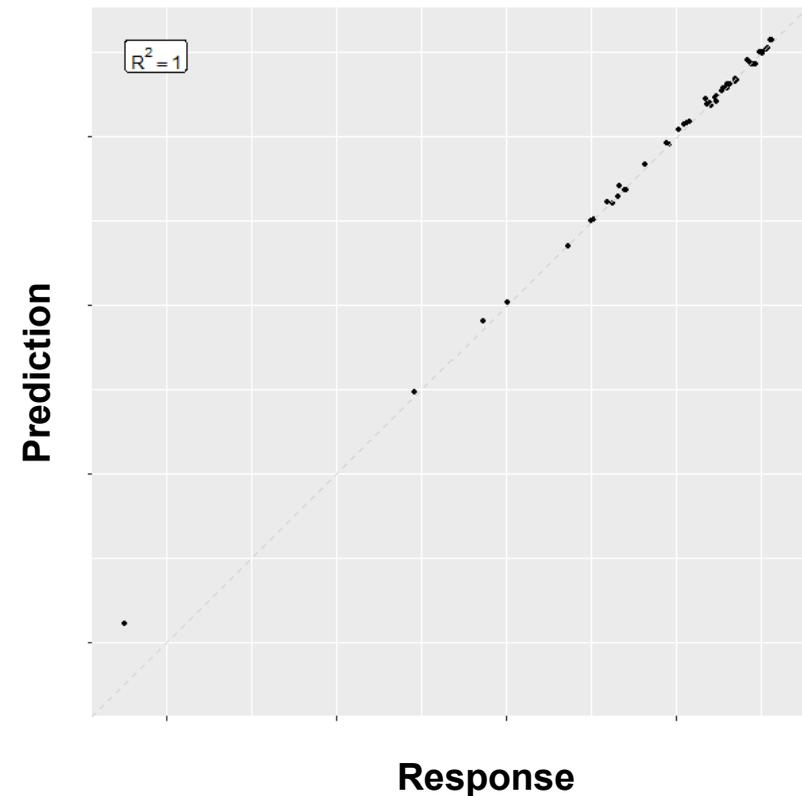
Aufgabenstellung

Beschreibung des PVFPs eines Lebensversicherers (Training auf 25`000 Szenarien) durch ein Neuronales Netz zu bestimmtem Stichtag

In-sample



Out-of-sample



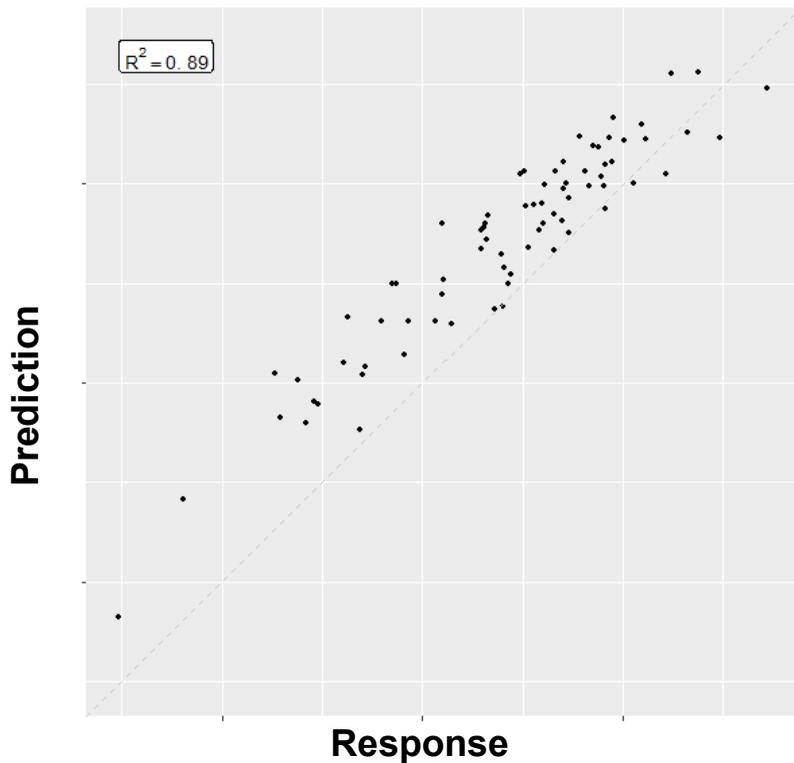
Erste Anwendungsbeispiele

Erklärung des PVFP durch graduelles Training

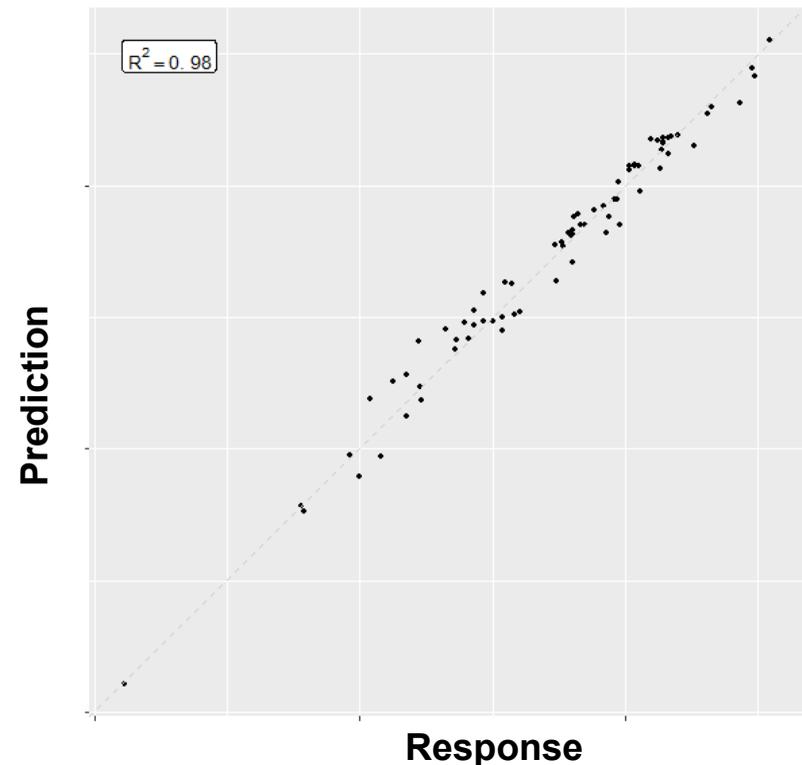
Run 2

Out-of-Sample Validierung desselben Neuronalen Netzes nach Übergang zum darauffolgenden Stichtag:

Out-of-sample: ohne Nachtrainieren



Out-of-sample: Nachtrainieren auf 5'000 Szenarien



Das graduelle Training ermöglicht eine effiziente Adaption des NN an (moderat) veränderte Zustände des Finanzmarkts und des Versicherungsbestands

Optimierung der SAA unter Solvency II Risikokennzahlen

Case Study

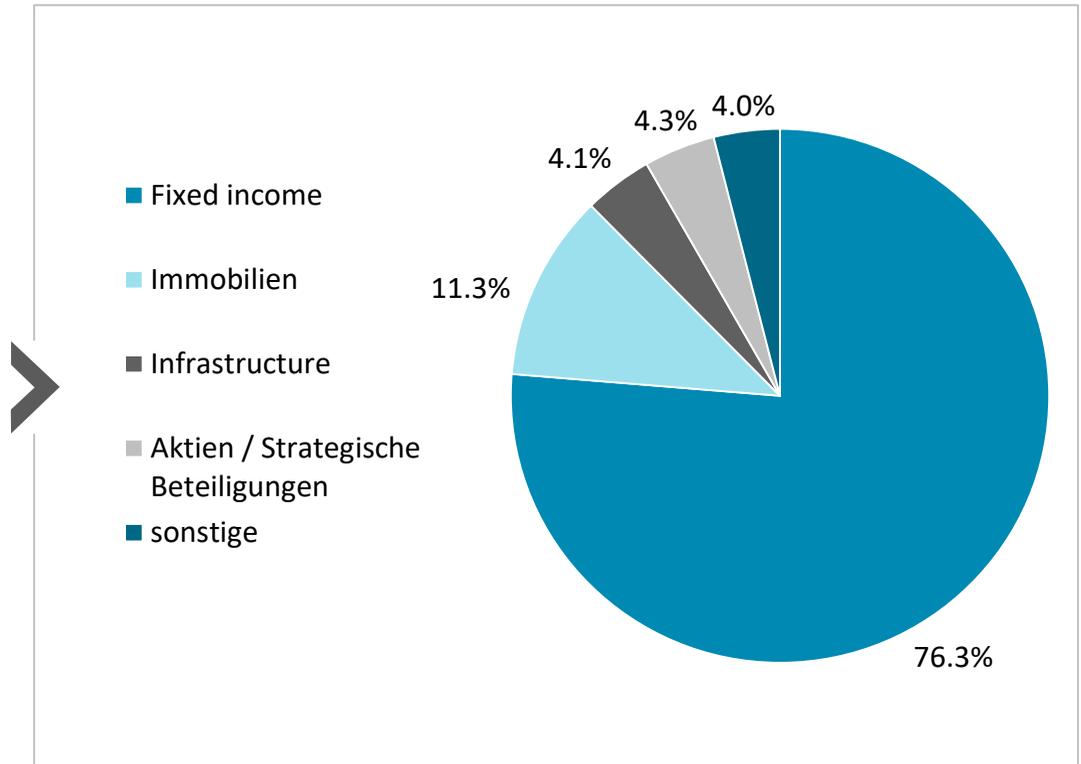
Hintergrund und Motivation

Motivation

Wie verändern sich in der klassischen Lebensversicherung die Own Funds und die Solvenzkapitalanforderung in Abhängigkeit von der Assetallokation?

Assetallokation

- Duration Assets: ca. 9
- Durchschnittsrating: A-
- SAA Ansatz:
 - Nachhaltig und dauerhaft das Garantierfordernis erwirtschaften
 - Zu jedem Zeitpunkt ausreichend Cash verfügen
 - Keine wesentliche Erhöhung des Risikos



Hintergrund und Motivation

Fragestellungen

Abschätzung der Wirkung typischer Maßnahmen der Strategischen Assetallokation in der Lebensversicherung:

1 Verringerung Durationsgap:

- Durationsgap nach Durationsverlängerung in verschiedenen Zinsszenarien?



2 Höhere Spreads und geringeres Durationsgap:

- Geringeres Gesamtrisiko bei besseren Erträgen?



3 Höhere Immobilienquote zu Lasten fixed income Portfolio:

- Sind Immobilien die bessere Assetklasse



4 Höhere Immobilien- und Aktienquote zu Lasten fixed income Portfolio:

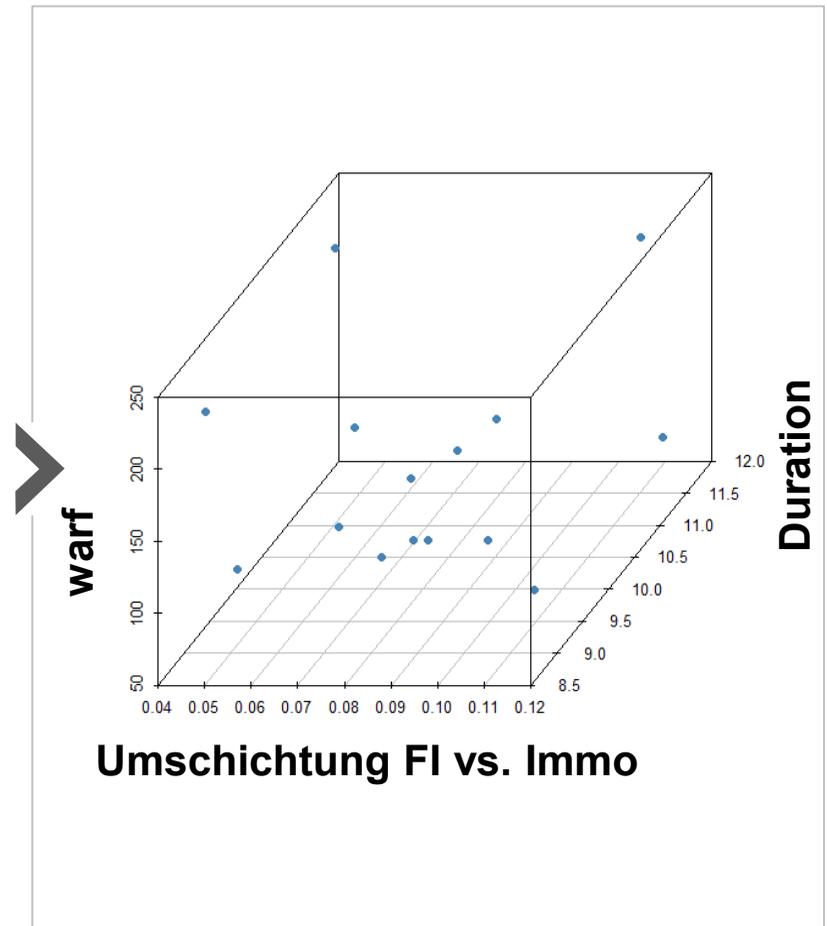
- Können Immobilienreserven das Risiko aus Aktien kompensieren



Vorgehen

Assetallokationen

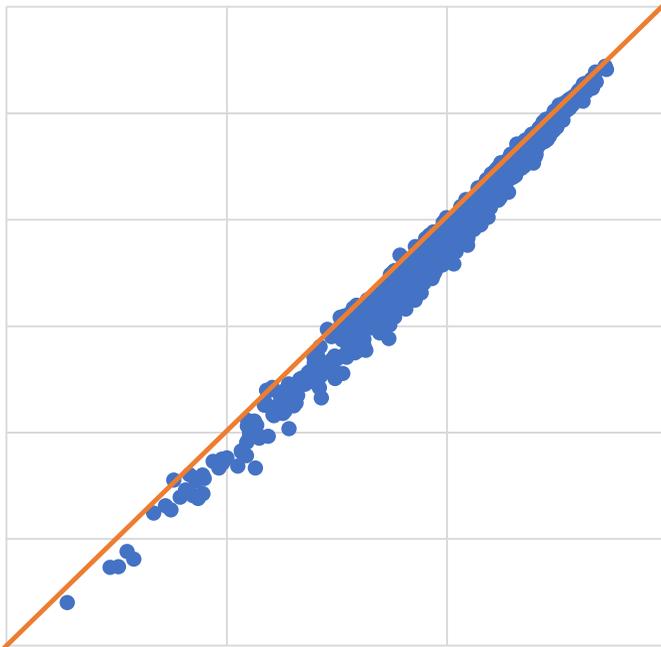
- Variationen auf drei Dimensionen
 - Allokation: Fixed Income vs Immobilien: von -5%- bis +5%-Punkte
 - Durchschnittliche Laufzeit: -2 bis +3 Jahre
 - Rating: BBB bis A+ (über warf)
 - 130 Kombinationen daraus
- Umschichtung proportional oder Reserve schonend
- Automatisierung (mittels python):
 - Umschichtung und neue Prophet Asset Tables
 - Neue Tables für Management Regeln
 - Parameter der Assetallokationen (Assetanteil, Duration, Warf, Reserven etc) für den Input in das neuronale Netz
 - Anpassung Marktwerte gemäß Stress des Risikofaktors (Market Value Adjustment)



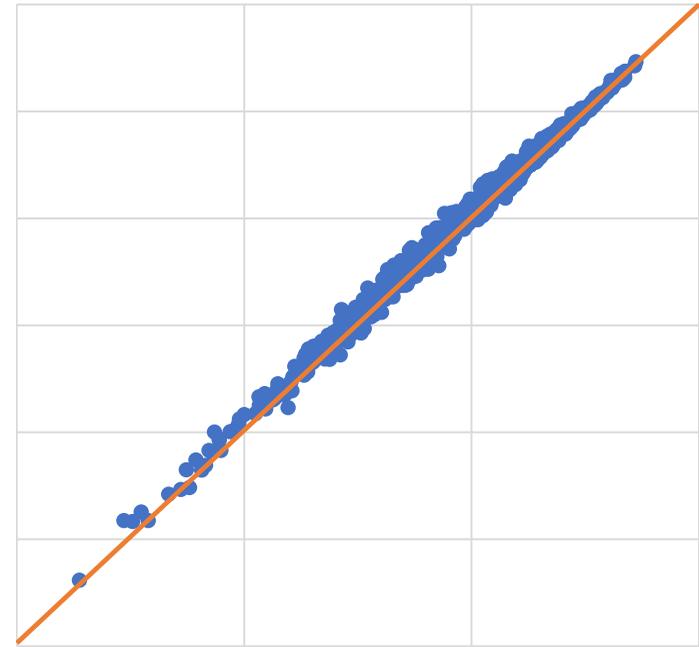
Vorgehen

PVFP aus Cashflow Modell in Abhängigkeit der Assetallokation

Basis vs. Allokation 90



Basis vs. Allokation 116

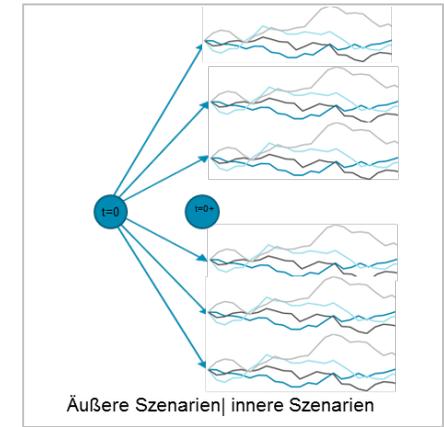


- Variation durch Assetallokation gering im Vergleich zu Effekt Szenarien
- Assetallokationen verändern vor allem Downside, kaum den Upside

Vorgehen

Szenarien

- 1'250 äußere Szenarien (Sobolsequenz) mit je 4 inneren Szenarien für Fitting, davon 1'125 für Training und 125 für Validierung
- Kombiniert mit je 15 Assetallokationen (Basisallokation + 14 Umschichtungen)
- 75 breit gestreute out-of-Sample Szenarien

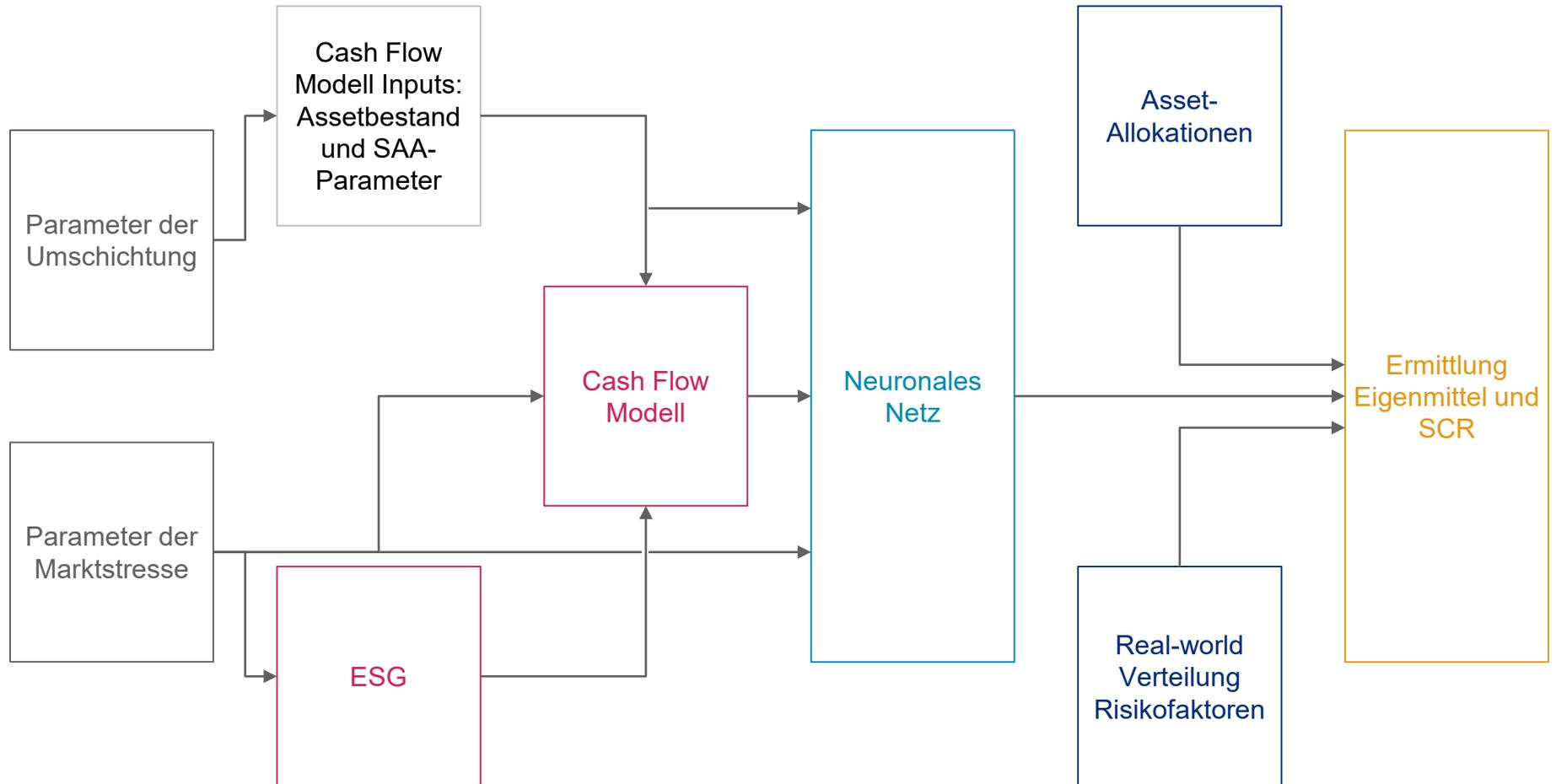


Neuronales Netz

- Umsetzung in KERAS
- Batchsize: 256, Epochen: 500 - 3'000, Aktivierungsfunktion: ReLU
- Anzahl getesteter Netze: ca. 30
- Beste Ergebnisse mit 5 Layern mit je 8 Neuronen
- Zielparameter: stochastischer PVFP / VIF
- Erklärende Variablen: Durationen und durchschnittliches Rating (Warn) von vier Anleiheklassen, relative (Marktwert-) Assetallokation, 17 Marktrisikofaktoren

Vorgehen

Prozess



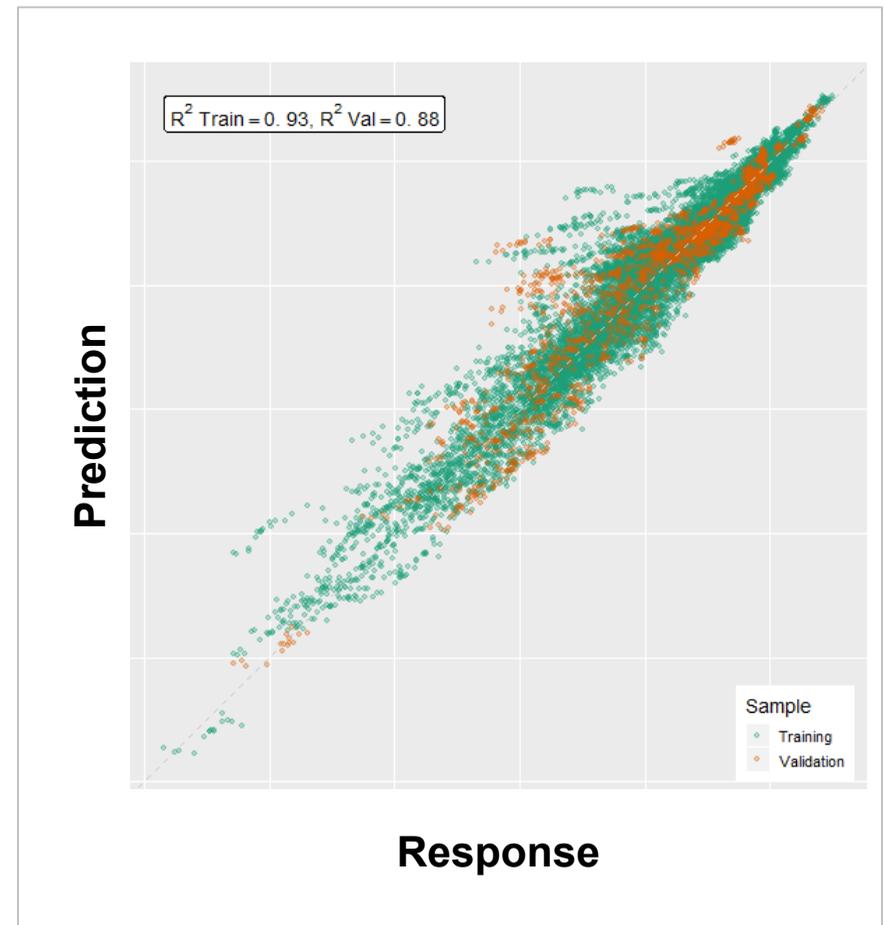
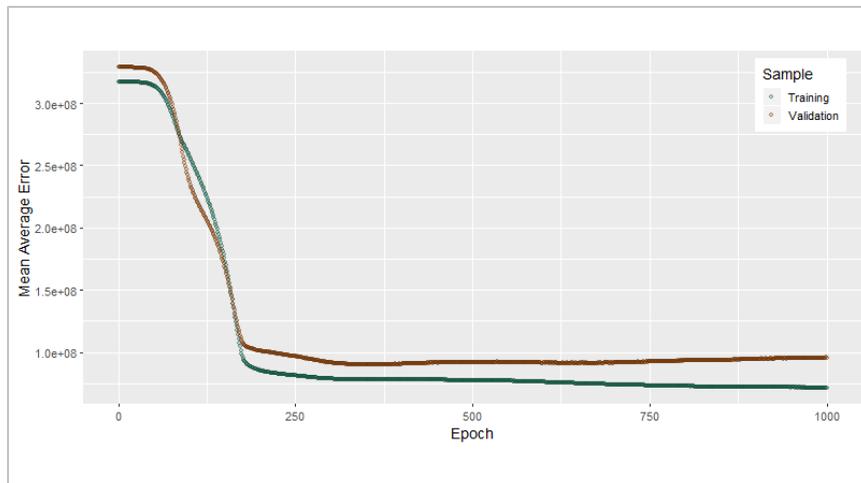
Trainingsergebnisse

Goodness-of-fit: In-sample

Beste Ergebnisse liefert ein NN mit 5 Layern à 8 Neuronen (trainiert für 1000 Epochen)

Beurteilung der Qualität:

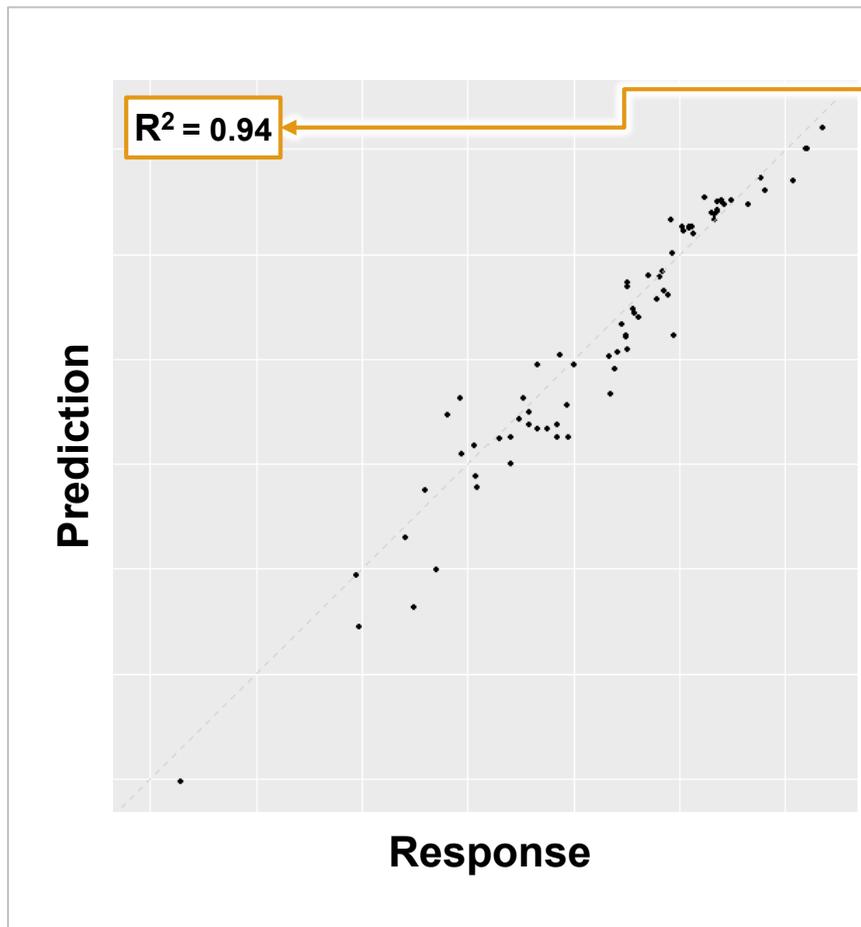
- Quantitativ: Abweichungen auf Trainings und Validierungsszenarien
- Qualitativ: Scatterplot In-sample, insb.:
Wie fügen sich die Validierungsszenarien in das Gesamtbild ein?
- Overfitting?



Trainingsergebnisse

Goodness-of-fit: Out-of-sample

Out-of-Sample Test nach fertiger Optimierung der Hyperparameter



Erklärungsgehalt (gemessen als R^2)
höher als auf den Trainingsszenarien

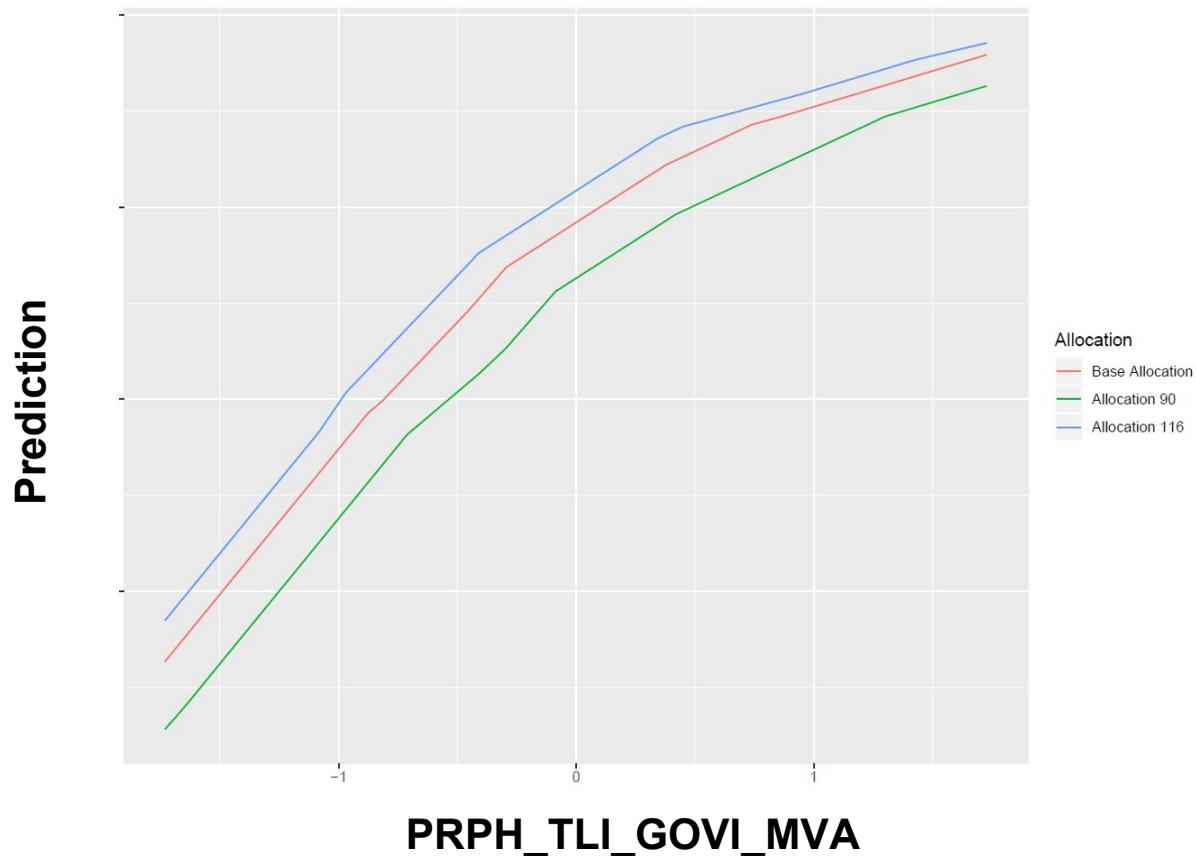
Out-of-Sample Validierung zeigt die finale Güte des NN bzgl.

- Interpolation
- Glättung
- » Verarbeitung von unbekanntem Inputsignalen/Risikofaktorausprägungen

Trainingsergebnisse

Univariate Proxyfunktionen

2d Plot der Auswertung des NN unter Variation der Änderung des initialen Marktwertes des Portfolios durch Government Bonds unter 3 verschiedenen Assetallokationen



Interpretation

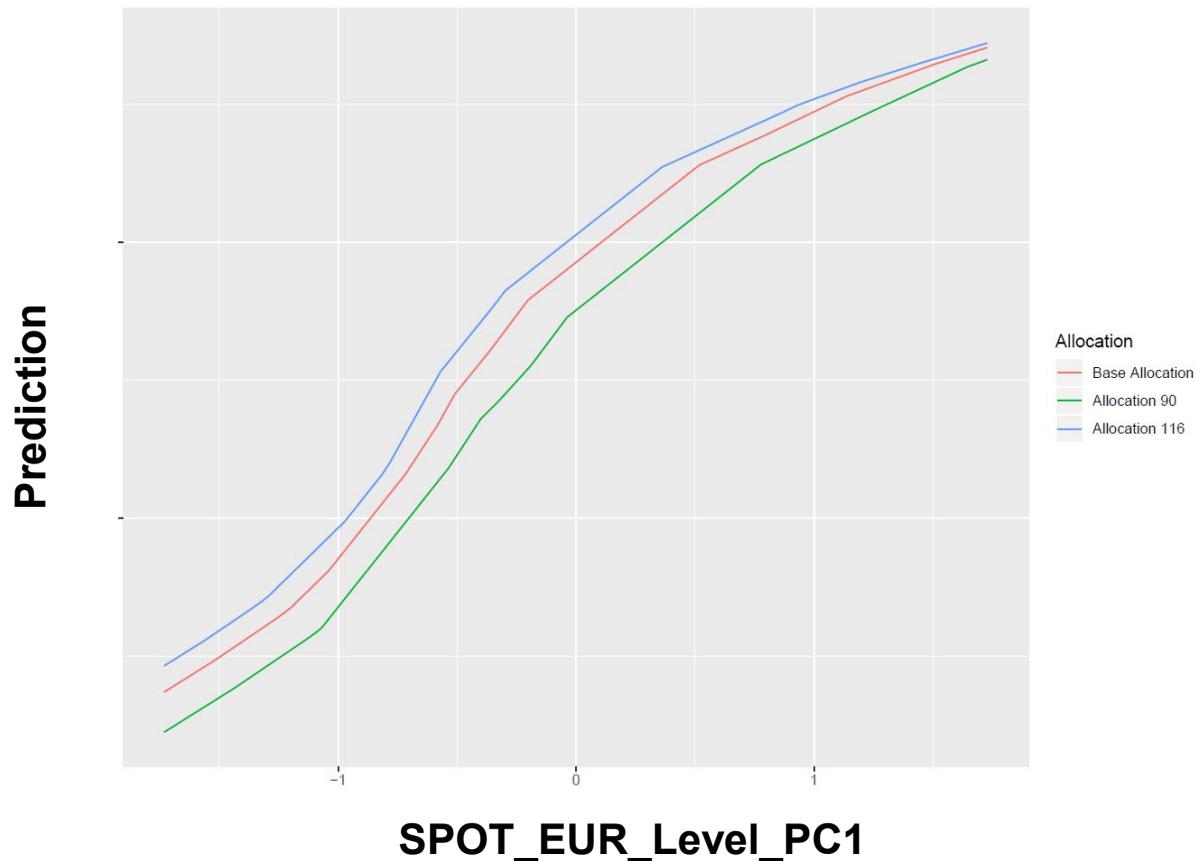


- Steigende Eigenmittel bei steigendem initialen Marktwert der Governmentbonds (=fallenden Spreads)
- Konkaver Funktionsverlauf spiegelt Asymmetrien des Geschäftsmodells wieder

Trainingsergebnisse

Univariate Proxyfunktionen

2d Plot der Auswertung des NN unter Variation der Änderung des initialen Zinsniveaus (parametrisiert über erste Hauptkomponente) unter 3 verschiedenen Assetallokationen



Interpretation

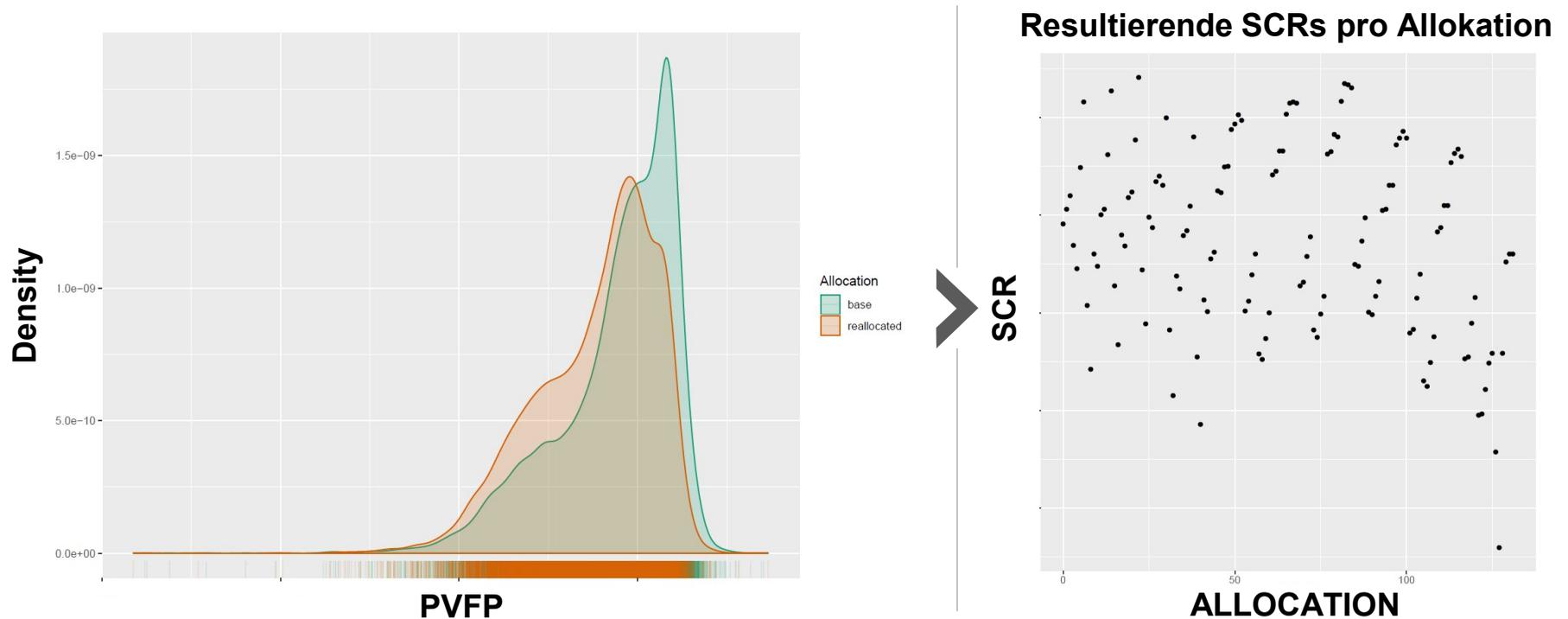


- Steigende Eigenmittel bei steigendem Zinslevel
- Konkaver Funktionsverlauf spiegelt Asymmetrien des Geschäftsmodells wieder

Anwendungen

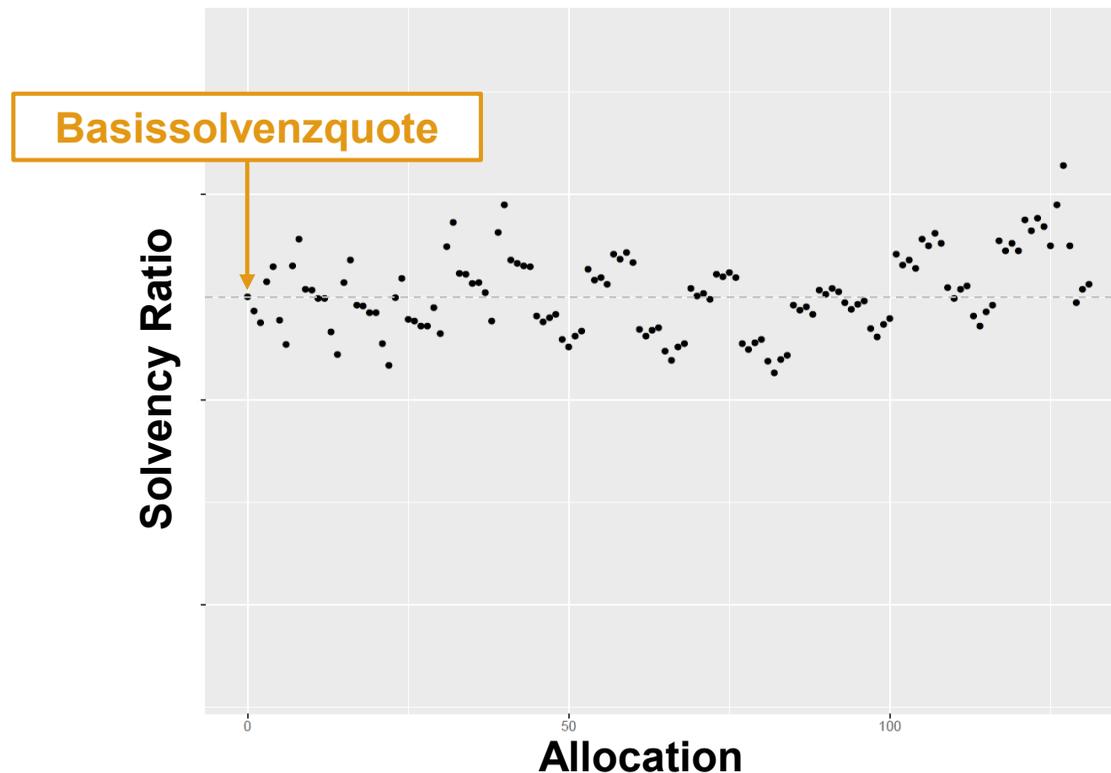
Verteilung der Eigenmittel pro Allokation

Die Auswertung des Neuronalen Netzes auf einer multivariaten Real-World Verteilung der Risikofaktoren liefert die Verteilung der Eigenmittel pro betrachteter Allokation (hier für 130 Beispielallokationen):



Das NN ermöglicht eine schnelle Beurteilung des Risikoprofils unter einer Vielzahl von Assetallokationen

Resultierende Solvenzquoten pro betrachteter Allokation

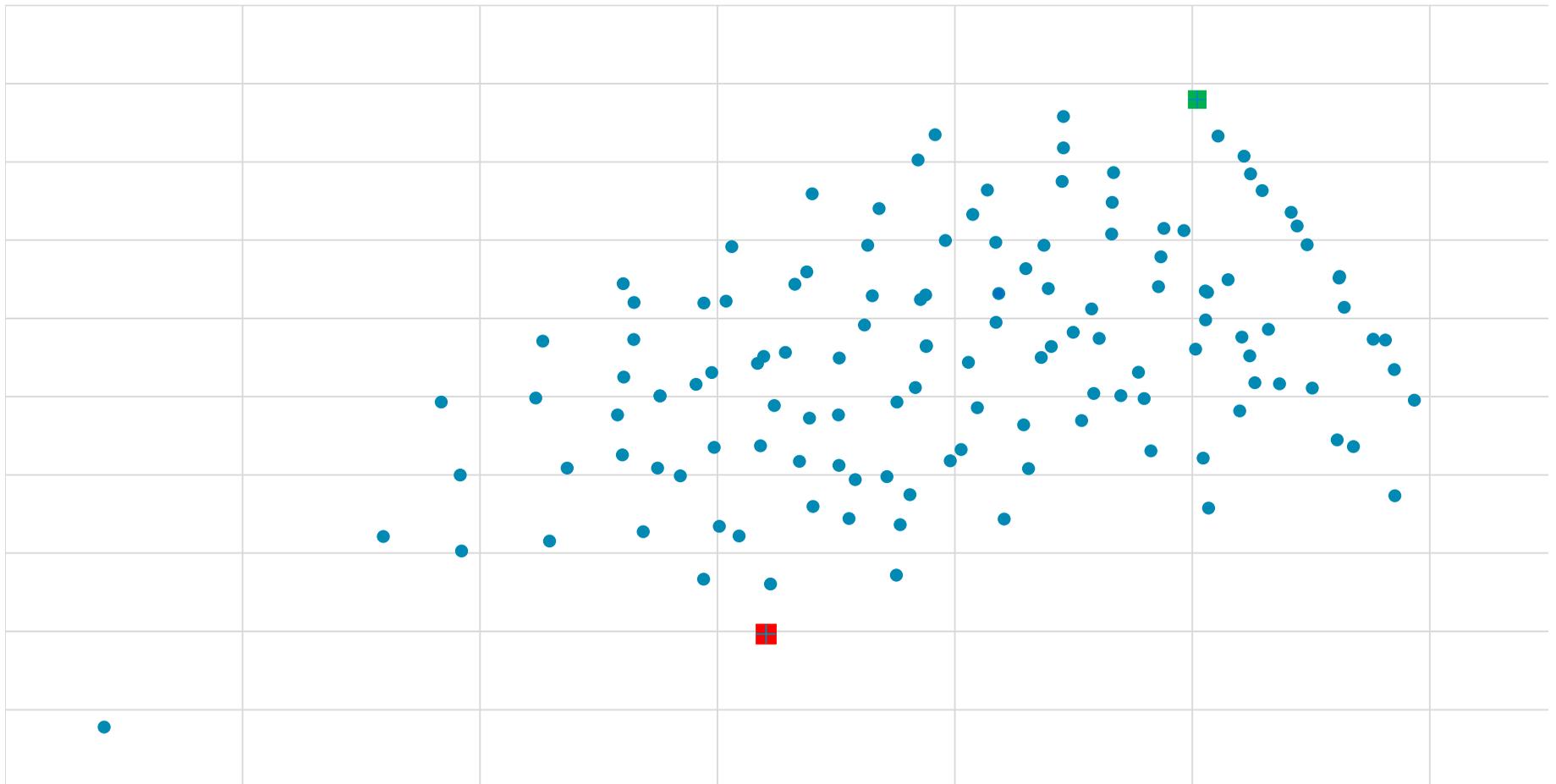


Das NN ermöglicht eine schnelle Beurteilung des Risikoprofils unter einer Vielzahl von Assetallokationen

Anwendungen

Maximierung der Eigenmittel unter Sicht des Risikokapitals

SCR vs PVFP pro Allokation auf einer real-world Auswertung von 130 Allokationen

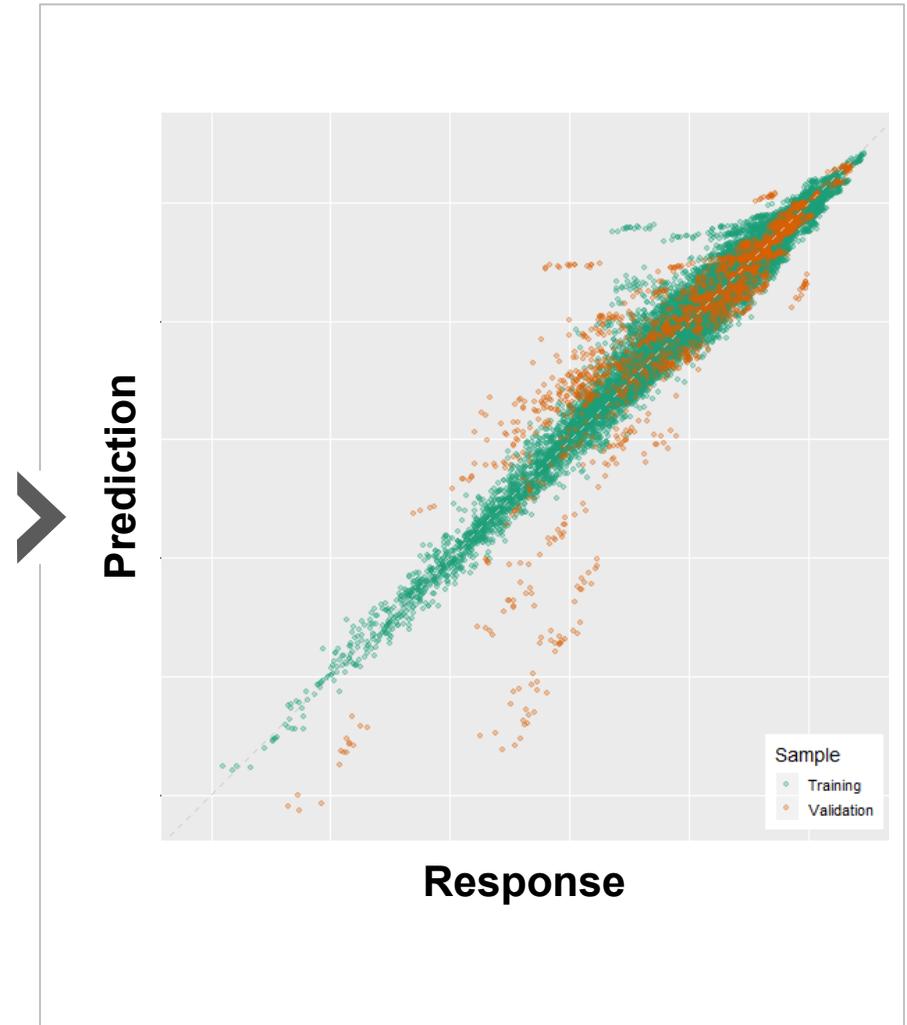


Erkenntnisse

und Ausblick

Weitere Erkenntnisse

- Daten, Daten und Prozesse:
 - Hauptaufwand in Datengewinnung und Verarbeitung
 - KERAS sehr effizient: testen verschiedener Varianten einfach möglich
 - Für eine effiziente Nutzung hoher Automatisierungsgrad notwendig
- Neuronal Netze sind sehr sensitiv gegen Overfitting
- Gute / deutliche Hinweise aus Ergebnissen des Neuronal Netzes auf
 - Unzulänglichkeiten des Cashflow Modells bzw. der Szenarien
 - Der Wirkung der Managementregeln im Cashflow Modell



Ausblick

- 1 Schätzung Own Funds über Umstellung auf (Assets minus Liabilities)

- 2 Weitere Risikofaktoren (separat) berücksichtigen: z.B.
 - Reserven (Reservehebungen)
 - Managementregeln
 - (Kennzahlen der) Deckungsrückstellung um z.B. Ablauf /Alterung zu berücksichtigen
 - Volatilität der (Buchwert) Veränderungen

- 3 Weitere Zielwerte: z.B. Verlauf Jahresüberschuss lokale Rechnungslegung

- 4 Einbindung in produktiven Prozess, Automatisierung

- 5 Berechnung auf real world Szenarien mit Drift für Analyse der economic Earnings in Abhängigkeit ECR

- 6 Synthetische Portfolien zur Analyse von Einzeleffekten (z.B. Zinsniveau im Verhältnis zur Garantie)

- 8 Weitere Verfeinerung neuronales Netz