

# VALIDIERUNGSPROZESS FÜR SCORING- UND RATING MODELLE MITTELS NEURONALER NETZE

Dr. Anatoliy Antonov, Dr. Ventsislav Nikolov  
09. Januar 2017

Eurorisk Systems Ltd.  
31, General Kiselov str.  
9002 Varna, Bulgaria  
Phone +359 52 612 367  
Fax +359 52 612 371  
[info@eurorisksystems.com](mailto:info@eurorisksystems.com)  
[www.eurorisksystems.com](http://www.eurorisksystems.com)

**Zusammenfassung:** *Diese Forschungsarbeit untersucht die Validierung und Kalibrierung von Modellen, die Kredit Scoring und Rating mit statistischen Methoden bestimmen. Dies wird erreicht, indem die Ergebnisse dieses Modells mit den Ergebnissen eines alternativen Modells, das auf einem neuronalen Netzwerk basiert, sowie die Berechnungen verschiedener statistischer Parameter verglichen werden. Vorgestellt wird ein Prototyp des Softwaresystems zur Analyse und Auswertung, der Distanz, Standardabweichung, und Korrelation, sowie des kumulatives Genauigkeitsprofil (CAP). Außerdem berechnet es die Akkumulation und Analyse historischer Statistiken für Ausfallausfälle.*

**Key words:** *Kredit-Rating, Scoring, Validierung, Analyse, Kalibrierung, neuronale Netze.*

Die Validierung von Modellen repräsentiert einen iterativen Prozess der in der Verantwortlichkeit der Bank liegt. Dieser Prozess umfasst bestimmte objektive quantitative sowie subjektive qualitative Faktoren, zusammen mit der Kalibrierung von gewichteten Koeffizienten, Sättigungsgraden, obligatorischen Bedingungen (K.O.-Prüfungen) und der endgültigen Klassifizierung der Scoring- und Rating Ebenen.

Software-Hersteller und Software-Implementierer können die Regulierungsbehörden bei der Vorbereitung und Durchführung von Modellvalidierungen unterstützen. Dies erfolgt mittels Analysen, Vorbereitungen von Beispielen und Teilnahmen an Diskussionen über Validierung und Kalibrierung. Die Modelle werden anhand statistischer Methoden und regressiver Prozeduren, sowie neuronaler Netze und CAP-Berechnungen (CAP = Cumulative Accuracy Profile und Gini Koeffizient) durchgeführt. In der Literatur existieren zahlreiche Analysen und Studien zur Modellvalidierung (siehe [www.google.de](http://www.google.de): Basel III rating system validation), die zur Erfüllung der Anforderungen von Aufsichtsbehörden verwendet werden können.

This research paper introduces specific principles and examples for the validation of models using neural networks. Eurorisk Systems Ltd. has a longstanding experience in using neural networks in various financial areas, such as clustering and prediction of factors, identification of distributions, calculation of implied ratings and multifactor models, which are implemented in different modules within the software systems of financial institutions.

In dieser Forschungsarbeit werden einige spezifische Prinzipien und Beispiele zur Validierung mittels neuronaler Netze dargestellt. Eurorisk Systems Ltd. verfügt über langjährige Erfahrung in der Nutzung neuronaler Netze in diversen Finanzbereichen, wie z. B. bei Clustering und Voraussage von Faktoren, Erkennung von Verteilungen, Berechnung von implizierten Rating und multifaktoriellen Modellen, die in verschiedenen Modulen innerhalb der Softwaresysteme von Finanzinstituten implementiert sind.

## 1. Das Objekt der Validierung

### **Ausgangssituation:**

Gegeben ist eine Vielzahl von Kreditnehmern mit deren Daten, wie z. B. Alter, Ersparnisse, Einkommen, Immobilien, usw. Für jeden Kreditnehmer wird eine Kredit Scoring bzw. Rating gegeben, das in der Vergangenheit anhand des zu validierenden Rating-Systems vergeben wurde.

## Ziel:

Das Ziel besteht darin, ein mathematisches Alternativmodell zur Validierung zu erstellen, und zwar anhand vorhandener Daten des Kreditnehmers und dessen Scoring / Rating-Werte zu, so dass:

- Bei neuen Kreditnehmerdaten, dessen Scoring / Rating noch unbekannt ist, kann das Modell den Wert des Scorings / Ratings bestimmen und diesen mit dem Ergebnis des Ratingsystems vergleichen.
- Ein Vergleich zwischen Scoring- und Rating-Werten ist möglich durch die Verwendung von Ratingsystemen und mathematischen Modellen nach Kalibrierung, bzw. nach dem Training des mathematischen Modells. Ziel ist die Validierung und Berechnung statistischer Daten, wie z. B. Fehlermarge, Standardabweichung, Korrelation, CAP, Gini-Koeffizient usw.

## 2. Neuronales Netz

Das definierte mathematische Modell basiert auf einem mehrschichtigen neuronalen Netzwerk, das mit symbolischen und numerischen Daten (Faktoren) eines Kreditnehmers arbeitet. Diese Daten wurden im Vorfeld in geeigneter Weise umgewandelt und verarbeitet. Die im neuronalen Netzwerk erwähnten Schichten sind wie folgt:

- Eingangsschicht: die Zahl der Neuronen = Zahl der Eingangsfaktoren, z. B. 24
- Versteckte Schicht: die Zahl der Neuronen wird nach Regeln bestimmt, z. B. 10
- Ausgangsschicht: die Zahl der Neuronen verbindet sich mit der Zahl der Ergebnisse, z. B. 2 Neuronen für Scoring und Rating

Die Struktur des Netzwerks ist in Abb. 1 dargestellt. Die Neuronen sind mit gewichteten Verbindungen  $W_{ij}$  und  $W_{kj}$  untereinander verbunden, die während des Trainings automatisch angepasst werden. Auf diese Weise realisiert das neurale Netzwerk ein hierarchisches Scoring auf zwei Ebenen. Die Neuronen werden über eine Polynomfunktion modelliert, an der die Ausgangssignale der vorhergehenden Neuronen beteiligt sind. Danach wird eine interne nichtlineare Aktivierungsfunktion angewendet, die den Neuronenausgang erzeugt.

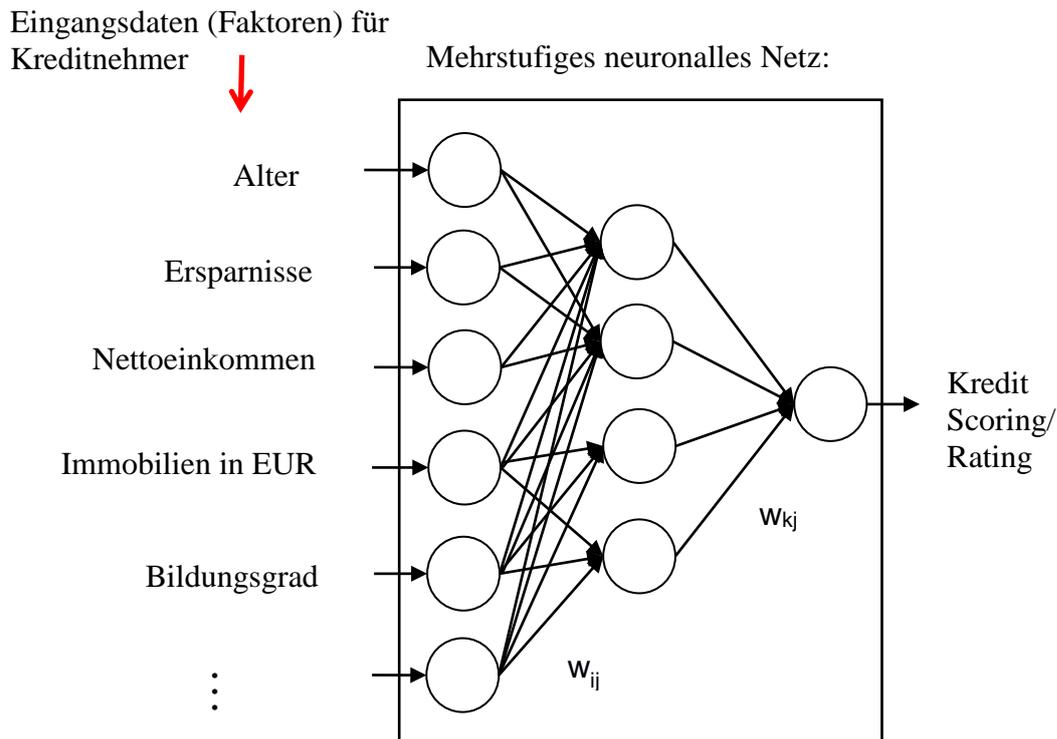


Abb. 1. Struktur des neuronalen Netzes

Das neuronale Netzwerk arbeitet in zwei Phasen:

### Stufe 1: Training des Netzes

Ein Satz von Eingangsdaten (Faktoren) jedes Kreditnehmers assoziiert sich mit den, im Vorfeld durch das Scoring / Rating-System bestimmten Scoring / Ratings. Anhand der Eingangsdaten wird die innere Struktur des neuronalen Netzes so eingestellt, dass die Abhängigkeiten zwischen den Eingangsfaktoren und den Ausgangsergebnissen durch die Einstellung der gewichteten Verbindungen „lernt“. Diese Stufe zeigt Abb. 2. Das Training erfolgt automatisch, in mehreren Epochen (aufeinanderfolgenden Trainingsitzungen).

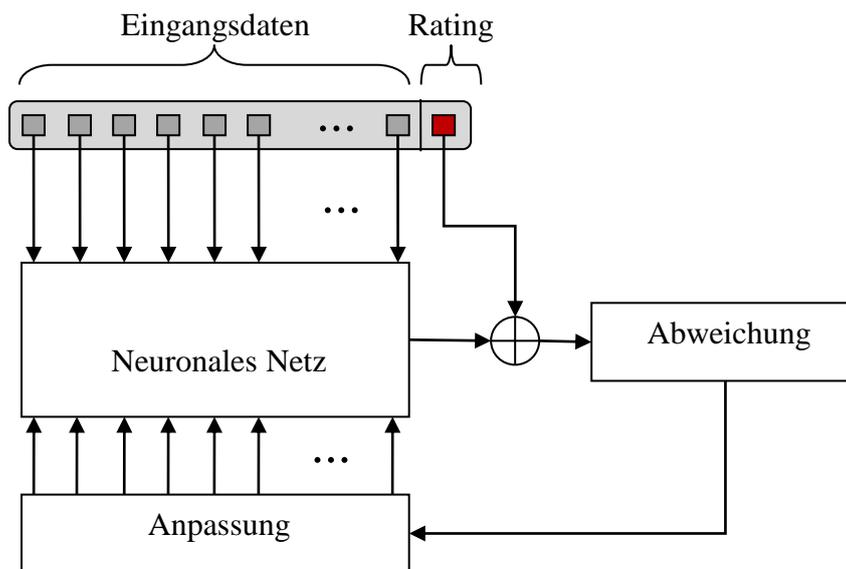


Abb. 2. Training des neuronalen Netzes

## Stufe 2: Benutzung des Netzwerkes

Neue Eingabedaten werden dem im Vorfeld trainierten neuronalen Netz hinzugefügt, wodurch der Wert des unbekannten Kredit-Scorings bzw. Ratings erzeugt wird.

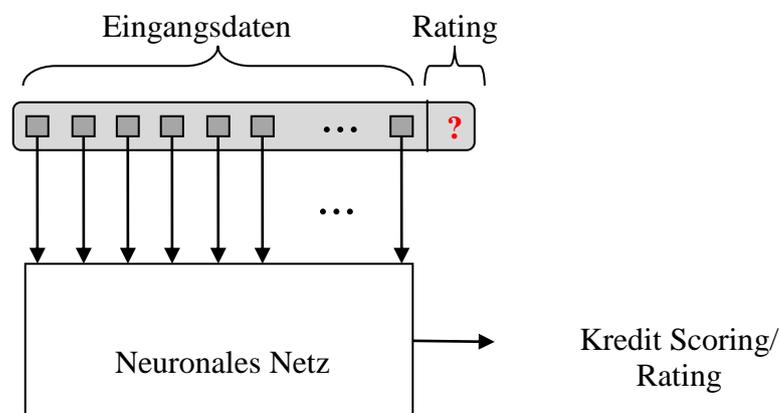


Abb. 3. Definition des Kredit-Scorings bzw. -Ratings mittels trainierter neuronaler Netze

### 3. Prototypmodul

Die Arbeit des neuronalen Netzes wird durch einen Prototyp veranschaulicht, welches auf echten historischen Daten basiert und die Validierungsaufgabe realisiert (siehe oben 1. Das Objekt der Validierung).

Ein Satz von 24 Eingabedaten (Faktoren) für 821 Kreditnehmer (gruppiert je nach Kredittyp in 3 Gruppen: Hypothekendarlehen, Verbraucherkredit oder Overdraft) werden in das neuronale Netzwerk geladen (Abb. 4). Als erstes wird ein Kredittyp ausgewählt. Nachdem das Netzwerk konfiguriert wurde, wird es trainiert, wobei alle Kreditnehmer in der Gruppe berücksichtigt werden. Danach sind individuelle oder "batch" Auswertungen möglich.

Das Validierungsprinzip basiert auf der Gegenüberstellung jedes einzelnen Scorings des Ratingmodell zu jedem einzelnen Scoring des trainierten neuronalen Netzwerks. Das akkumulierte Wissen des Netzes entspricht der Erfahrung aller Scoringvorgänge während des Trainings.

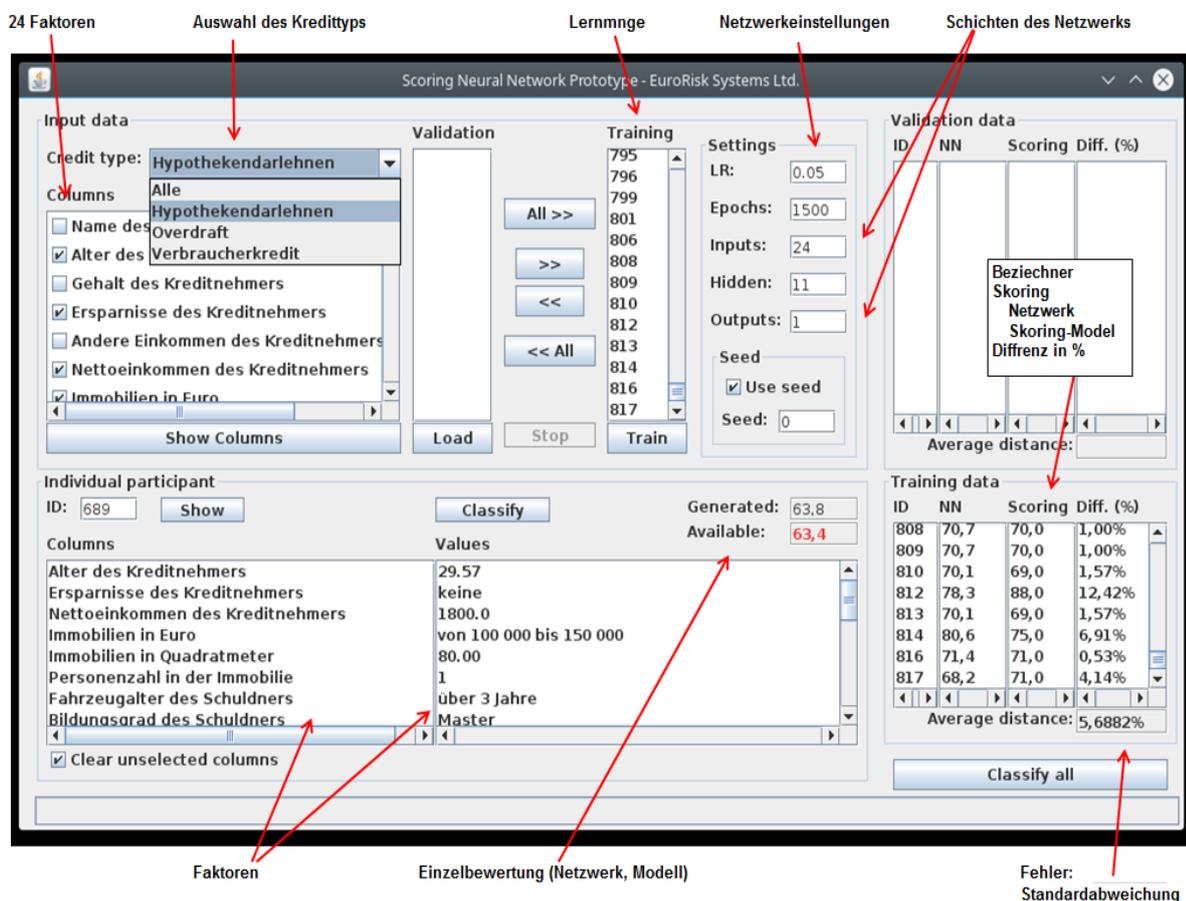


Abb. 4. Prototyp des neuronalen Netzes. Training und Validierung für Hypothekendarlehen

Das Ergebnis dieser Gegenüberstellung wird statistisch bearbeitet und kann zur Bewertung und Validierung des Scoring- und Rating Modelle genutzt werden, z. B. Fehler, Standardabweichung, Korrelation und CAP. Die Statistiken zeigen, inwiefern die Regeln zum Scoring in dem Scoring-Modell konsistent in Bezug auf Zusammenfassung sind und eine gemeinsame Strategie entsprechen oder inwiefern sie auf mehreren Ausnahmen und individuellen Entscheidungen basieren.

Nummer	Factor	Beispiel: Hypothekendarlehen	Beispiel: Overdraft
1	Alter	51,62	47,38
2	Ersparnisse	1 000 - 10 000	keine
3	Nettoeinkommen	1434	0
4	Immobilien in Euro	über 150 000	über 150 000
5	Immobilien in Quadratmeter	72	200
6	Zahl der Personen im Haushalt	3	1
7	Fahrzeugalter	über 3 Jahre	über 3 Jahre
8	Bildungsgrad	Abitur	Master
9	Familienstand	verheiratet	verheiratet
10	Arbeitgeberrechtsform	AG	Mitgliedsunternehmen
11	Branche	Sonstiges	Sonstiges
12	Tätigkeitsbezeichnung / Angestellter in	Lokales Unternehmen	Lokales Unternehmen
13	Einkommensquelle	Unbefristete Festanstellung	UnbefristeteFestanstellung
14	Berufserfahrung in Jahren	25	23
15	Untertyp	Bürger	Bürger
16	Gemeinde	Berlin	Hamburg
17	Zielgruppe der Bank	Nein	Zielgruppe I
18	Kredittyp	Hypothekendarlehen	Overdraft
19	Höhe des Kredits	61360	1500
20	Zinssatz	8,15	10,95
21	Währung	EUR	BGN
22	Kreditrückzahlung in Monate	216	24
23	Tilgungsart	Annuitätenplan	
24	Methode zur Einkommensprüfung	Geprüftes Einkommen	
	<b>Scoring</b>	<b>72,80</b>	<b>58,40</b>

Abb. 5. Faktoren für Scoring / Rating und zwei Beispiele

Die Tabelle aus Abb. 5 zeigt eine Liste mit Hauptfaktoren von zwei verschiedenen Kreditnehmern. Die Daten des Kreditnehmer werden numerisch bearbeitet, bevor sie dem neuronalen Netz hinzugefügt werden. Die Daten in der letzte Zeile zeigen die Scoring-Ergebnisse, die mit dem Scoring-Modell berechnet wurden. Scoring-Werte liegen zwischen 0% und 100%. Daten und Scoring-Ergebnisse aller Kreditnehmer aus der ausgewählten Gruppe werden für das Training des Netzwerks verwendet.

Mit Hilfe des Prototyps in Abb.5 wird ein Training mit nur einem Teil der Daten und der Scoring-Ergebnisse der Kreditnehmer (trainierende Menge) durchgeführt. Der Rest (validierende Menge) erscheint dem neuronalen Netzwerk als unbekannt. In diesem Fall zeigen die Bewertungsergebnisse der Validierung an, in welchem Ausmaß das Netzwerk mit unbekanntem Daten umgehen kann, d.h. wie leistungsfähig die Diskriminanzfunktion des Scoring-Modells ist.

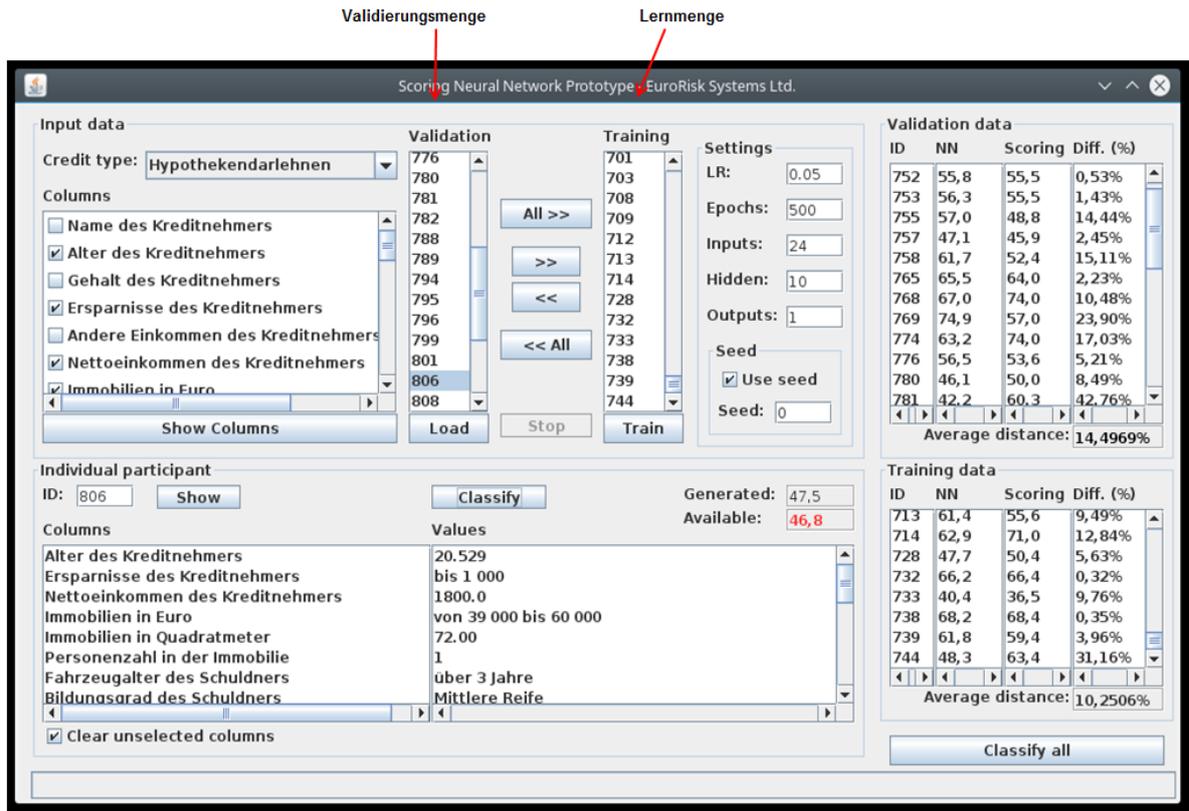


Abb. 6. Prototyp des neuronalen Netzes: Training und Scoring neuer Hypothekendarlehen

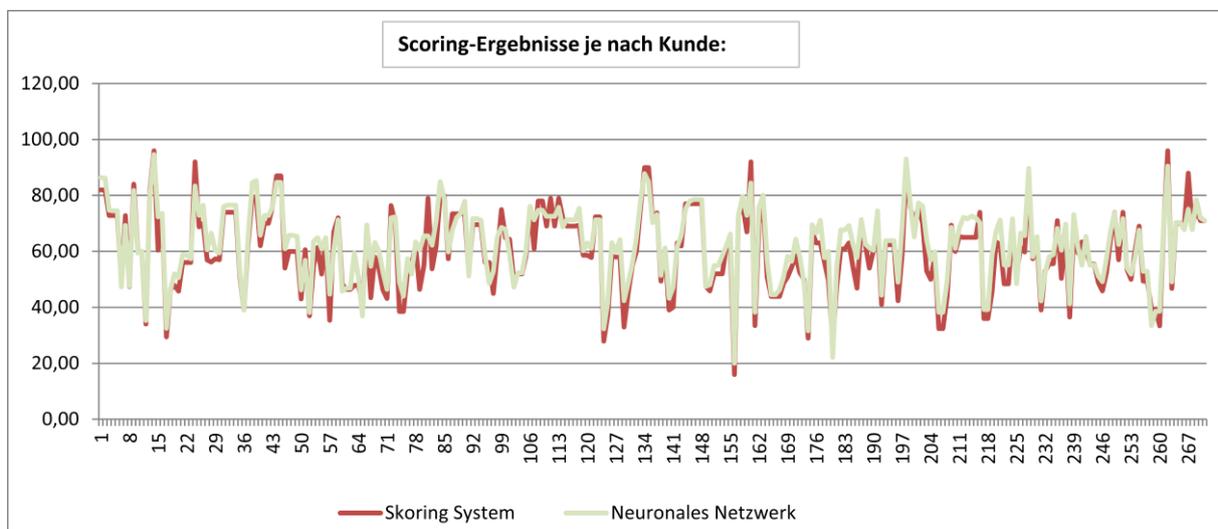
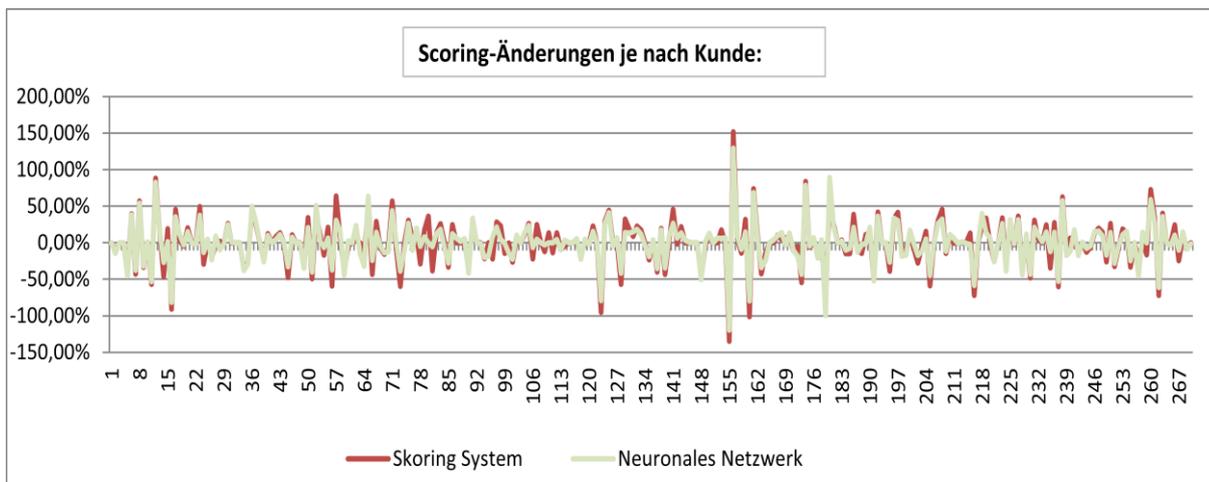


Abb.7. Scoringssystem versus neuronales Netz

Abb. 7 stellt einen graphischen Vergleich der Scoring-Ergebnisse aus dem Scoring-Modell und der Scoring-Ergebnissen aus dem trainierten neuronalen Netzwerk für die Hypothekendarlehensgruppe dar. Es gibt eine gewisse "Nüchternheit und Ausgewogenheit" des Netzwerks, die nicht auf geringfügige Änderungen im Scoring-Modell reagiert. Dieses Verhalten wird erwartet, da das Scoring des Netzwerks auf dem in der Trainingsphase erworbenen "Wissen" basiert.

Die relativen Änderungen sind in Abb. 8 dargestellt, wo ein solches Verhalten auch beobachtet werden kann.



*Abb 8. Scoringssystem versus neuronales Netz*

In der Abb. 9 ist der Vergleich der Scoring-Ergebnisse der ersten 29 Kreditnehmer von Hypothekendarlehen (insgesamt 271 Kreditnehmer) abgebildet. Die Statistik zeigt eine hohe positive Korrelation von 0,9431 zwischen den Ergebnissen des Scoring-Modells und denen aus dem neuronalen Netzwerk. Die Standardabweichung von beiden Scoring-Ergebnissen beträgt 5,27%, was auf eine gute Übereinstimmung und Stabilität des Scoring-Modells hinweist. Aus diesen statistischen Ergebnissen folgt die Stabilität und die Validierung des Scoring-Modells.

#### **Anmerkungen:**

- Die Validierung der Scoring- und Rating-Modelle kann auch anhand von historischen Ausfallstatistiken durchgeführt werden. Dies erfolgt unter der Voraussetzung, dass die Daten aus einem größeren Satz historischer Daten gespeichert, synchronisiert und akkumuliert wurden. Wenn dies der Fall ist, können Rating-Systeme anhand historischer Daten validiert und kalibriert werden.
- In der o.g. Methodologie ist die Validierung autoregressiv, d.h. dass eine Reihe von Scoring-Ergebnissen aus den Scoring-Modellen selbst das "Scoring-Kenntnisse" bestimmt.
- Die wichtigsten Faktoren können identifiziert werden, indem einzelne Faktoren ausgeschlossen werden und die Veränderung der GAP (CAP = Cumulative Accuracy Profile) bewertet wird, die mit der Gewichtung des ausgeschlossenen Faktors vergleichbar ist.

Training		Standartabweichung				5,27%
Mittelwert	62,67	60,26	Korrelation		0,9431	
Kundennummer	Scoring-Ergebnisse			Logarithmische Änderung		
	Neuronales Netzwerk	Scoring-System	Differenz in %	Neuronales Netzwerk	Scoring-System	
1	86,40	82,00	5,13%	-0,23%	0,00%	
2	86,20	82,00	4,85%	-14,32%	-11,90%	
3	74,70	72,80	2,52%	-0,27%	0,00%	
4	74,50	72,80	2,26%	0,00%	0,00%	
5	74,50	72,80	2,26%	-45,43%	-39,59%	
12	47,30	49,00	3,64%	38,19%	39,59%	
16	69,30	72,80	5,04%	-37,56%	-42,91%	
21	47,60	47,40	0,40%	54,14%	57,22%	
23	81,80	84,00	2,68%	-32,00%	-33,65%	
26	59,40	60,00	1,01%	1,17%	0,00%	
27	60,10	60,00	0,25%	-53,21%	-56,80%	
32	35,30	34,00	3,61%	82,56%	88,04%	
33	80,60	82,00	1,74%	16,02%	15,76%	
39	94,60	96,00	1,52%	-26,75%	-46,34%	
41	72,40	60,40	16,57%	1,51%	18,95%	
46	73,50	73,00	0,73%	-81,60%	-90,95%	
48	32,50	29,40	9,64%	35,39%	45,63%	
50	46,30	46,40	0,26%	11,61%	2,97%	
53	52,00	47,80	8,15%	-2,73%	-4,27%	
54	50,60	45,80	9,40%	15,19%	20,11%	
57	58,90	56,00	4,86%	-0,34%	0,00%	
58	58,70	56,00	4,61%	-2,24%	0,00%	
59	57,40	56,00	2,42%	37,36%	49,64%	
62	83,40	92,00	10,31%	-13,46%	-29,06%	
64	72,90	68,80	5,56%	4,82%	1,73%	
68	76,50	70,00	8,54%	-23,30%	-20,72%	
73	60,60	56,90	6,14%	9,29%	-1,24%	
76	66,50	56,20	15,53%	-9,79%	2,11%	
99	60,30	57,40	4,81%	-1,50%	-0,70%	
-----	-----	-----	-----	-----	-----	
271	Hypotheken-	darlehen	-----	-----	-----	

Abb. 9. Validierung von Hypothekendarlehen:  
Scoringssystem vs. neuronaler Netze

Die nächste Bewertung basiert auf eine CAP-Darstellung (Abb. 10), die das Scoring-Modell mit einem perfekten Modell vergleicht. Die Fläche unter der Kurve des Scoring-Modells wird auf der Fläche unter der Kurve bei dem perfekten Modell bezogen und somit das Gini Koeffizient =  $ap / (ap+ar)$  ermittelt. Gini bewegt sich zwischen 0% und 100%, wobei Scoring-Modelle mit höherer Qualität einen hohen Gini-Koeffizienten haben.

Der Entscheidungswert für das Scoring-Modell und das neuronale Netzwerk wurde auf 48% festgelegt. Kreditnehmer mit einem Scoring-Ergebnis von weniger als 48% werden in die Kategorie „Ausgefallen“ eingestuft. In dem Beispiel werden 36 Kreditnehmer bei dem neuronalen Netzwerk der Kategorie „Ausgefallen“ zugeordnet. Das Scoring-Modell erfasst 54 solcher Kreditnehmer.

Sowohl im Scoring-Modell als auch im neuronalen Netzwerk werden 25 Kreditnehmer in die Kategorie “Ausgefallen” eingestuft.

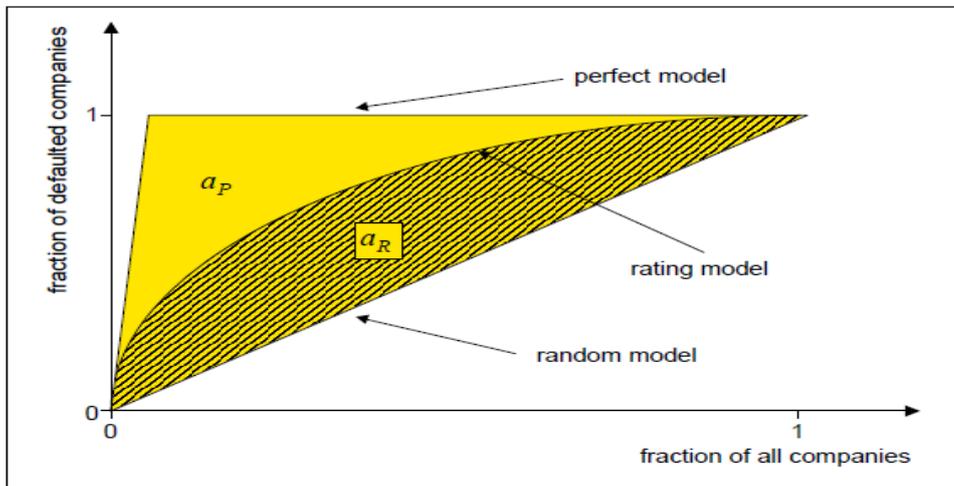


Fig. 10. CAP (Cumulative Accuracy Profile)

Abb. 11 ist die CAP-Darstellung mit den Scoring-Daten, die durch das Bewertungsmodell und das neuronale Netzwerk erhalten wurden. Ein Gini-Koeffizient von 80% kann abgelesen werden. Dabei wird angenommen, dass das durch das neuronale Netzwerk erzeugte Scoring den Erwartungen (Wahrscheinlichkeiten) für zukünftige Ausfälle entspricht.

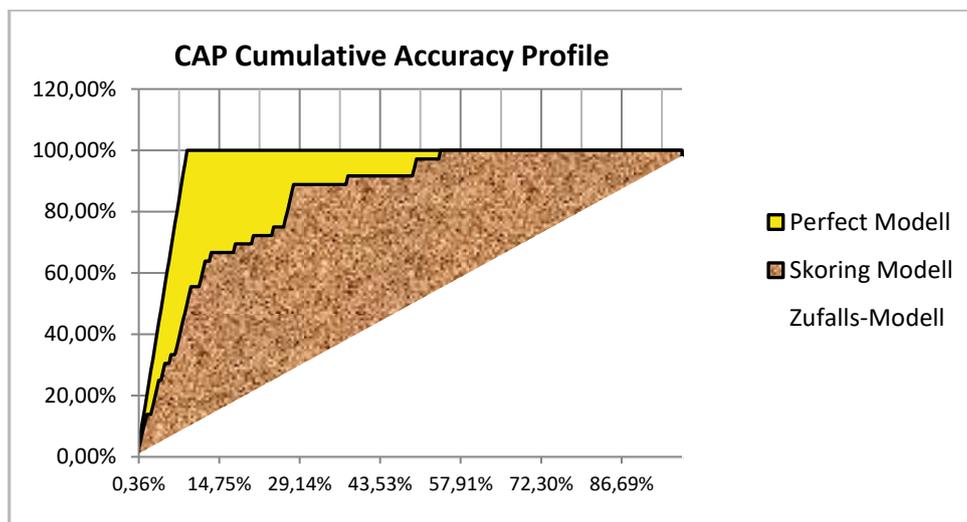


Abb. 11. CAP-Darstellung (Cumulative Accuracy Profile) nach dem Scoring