

## **4. cjelina: Kredit scoring modeli za stanovništvo (retail)**

- 4.1. Krediti za stanovništvo
- 4.2. Subjektivna kreditna analiza
- 4.3. Kredit scoring modeli za stanovništvo
- 4.4. Reference

#### 4.1. Krediti za stanovništvo

Kreditni stanovništvu obuhvaćaju kratkoročne i srednjoročne kredite koji se koriste za financiranje kupovine roba i usluga za osobnu potrošnju ili za ponovno financiranje dugova nastalih za takve svrhe. Postoje tri osnovna tipa potrošačkih kredita:<sup>1</sup>

##### 1. Krediti koji se otplaćuju u ratama

Podrazumijevaju periodično plaćanje glavnice i kamata. Obično potrošači pozajmljuju kako bi kupili neke trajne proizvode ili pokrili izvanredne troškove. Krediti koji se otplaćuju u ratama mogu biti direktni i indirektni. Direktni kredit se pregovara direktno između banke i korisnika koji traži kredit. Primjerice, osoba koja pozajmljuje od banke kako bi kupila automobil mora formalno zatražiti kredit i pružiti banci na uvid svoje osobne financijske informacije. Kreditni referent pregledava zahtjev za kreditom i donosi odluku o tome odobriti li kredit ili ne. Indirektan kredit financiran je od strane banke preko odvojenog trgovca koji prodaje robu potrošaču. U tom slučaju, trgovac automobilima dogovora uvjete s komitentom koje potom prezentira banci. Ako banka odobri takav kredit, komitent kupuje automobil prema prethodno dogovorenim uvjetima.

##### 2. Kreditne kartice i ostali revolving krediti

Banke nude različite kreditne kartice. Iako neke banke izdaju svoje kreditne kartice s vlastitim logom i podržavaju je vlastitim marketinškim naporima, većina ima franšizu za MasterCard ili Visa. Kako bi banka postala članom bilo kojeg od dva spomenuta sustava kartica, mora platiti članarinu koja se plaća samo jednom prilikom ućlanjenja te godišnji trošak koji je određen brojem klijenata koji aktivno koriste karticu. Za uzvrat MasterCard ili Visa preuzimaju nacionalni marketinški poduhvat. Kreditne kartice su atraktivne zato što osiguravaju povrate koji su prilagođeni riziku. Izdavatelji kreditnih kartica mogu zaraditi iz dva izvora: 1.) naplaćivanje godišnje članarine za vlasnike kreditnih kartica; 2.) naplaćivanje kamata na premašenu kreditnu razliku; Svaka kreditna kartica ima odobreni kreditni limit koji ograničava maksimalan dug koji se može prihvatiti u svakom trenutku. Kreditne kartice su toliko profitabilne

<sup>1</sup> Koch, T.W., MacDonald, S.S., Bank Management, Harcourt College Publishers, Fortworth, Tokyo, 2000, p. 701-709

---

zato što su mnogi potrošači cjenovno neosjetljivi. Naime, kamatne stope na plaćanje kredita kreditnim karticama su među najvišim kamatama koje se zaračunavaju uopće. Ono što je negativno kod kreditnih kartica je što su gubici među najvećima u svim tipovima kredita.

### 3. Krediti koji se ne otplaćuju u ratama

Ograničeni broj potrošačkih kredita zahtjeva plaćanje kamata i glavnice odjednom. Takav se kredit odobrava na temelju dobro definiranog budućeg tijeka gotovine. Takvi se krediti pojavljuju primjerice kada osoba posuđuje sredstva za polaganja avansa za novu kuću, a kredit se vraća u cijelosti, odjednom kada osoba proda postojeću kuću.

Krediti stanovništvu se međusobno toliko razlikuju da je nemoguće imati sveobuhvatan format analize koji bi se primjenjivao jednako na sve vrste kredita. Kreditna analiza se razlikuje s obzirom na različite tipove kredita stanovništvu. Karakteristike koje se tipično upotrebljavaju pri odobravanju kredita stanovništvu nalaze se u tablici 1. Treba naglasiti da te karakteristike variraju od situacije do situacije pošto je sasvim razumljivo da će se od zajmotražitelja koji traže kredite za kupovinu automobila tražiti drugačije informacije od onih koji traženim kreditom žele kupiti kuću. Kao što se vidi iz tablice 1., varijable su uglavnom kategorijalne, a i uobičajeno je da se kontinuirane kategoriziraju. Većina kredit scoring modela upotrebljava indikator varijable. No, međutim, alternativni pristup koji podrazumijeva upotrebu modela koji koriste kontinuirane varijable postaje sve uobičajenija.

**Tablica 1: Varijable koje se uobičajeno koriste pri odobravanju potrošačkih kredita**

Karakteristika (varijabla)
Vrijeme provedeno na jednoj adresi
Stambeni status
Poštanska oznaka
Telefon (ima, nema)
Prosječan godišnji prihod zajmotražitelja
Kreditne kartice (ima, nema)
Tip bankovnog računa (ček i/ili štednja, ništa)
Dob
Broj sudskih tužbi
Zanimanje
Namjena kredita
Bračni status
Koliko dugo ima račun u banci
Koliko dugo je na trenutnom poslu

Izvor: Hand, D.J., Henley, W.E., Statistical Classification Methods in Consumer Credit Scoring: a Review, Journal Royal Statistical Society A, 160, 1997, p. 523-541

#### 4.2. Subjektivna kreditna analiza

Cilj kreditne analize koja se provodi za odobravanje kredita stanovništvu je procijeniti rizik koji je povezan s odobravanjem kredita. Zbog toga što su fiksni troškovi obrade potrošačkih kredita vrlo visoki, banka mora generirati značajnu količinu takvih kredita kako bi smanjila troškove po jednom kreditu. To zapravo znači veliki broj različitih zajmotražitelja s različitim osobnostima i financijskim karakteristikama. Kao i kod kredita koji se odobravaju poduzećima, tako i kod potrošačkih kredita postoji 6 K<sup>2</sup>:

##### 1. Karakter

Karakter je najvažnije, ali isto tako i najteže procijeniti. Kreditni referent mora ustanoviti želju komitenta da vraća kredit. Jedina kvantitativna informacija koja je na raspolaganju jest aplikacija za kredit i kreditni registar. Ako je zajmotražitelj komitent banke, tada kreditni referent može provjeriti povijest njegova računa. Ako se pak radi o zajmotražitelju koji je komitent neke druge

<sup>2</sup> Koch, T.W., MacDonald, S.S., Bank Management, Harcourt College Publishers, Fortworth, Tokyo, 2000, p. 718-719

banke, tada se kreditni referent može obratiti kreditnom birou odakle će dobiti informacije o povijesti plaćanja za konkretnog komitenta. Banka se također u velikoj mjeri oslanja na subjektivnu ocjenu zajmotražitelja koju obavljaju kreditni referenti. Oni traže osobne reference, provjeravaju zaposlenje, točnost podataka iz zahtjeva za kreditom i slično, ne bi li ustanovili što bolje karakter komitenta.

## 2. Kapital

Kapital se odnosi na imućnost zajmotražitelja što svakako utječe na njegovu sposobnost vraćanja kredita.

## 3. Kapacitet

Kapacitet je financijska sposobnost komitenta da zadovolji otplatu kredita kao dodatak troškovima života i ostalim obvezama. Za gotovo sve potrošačke kredite prihod osobe je primarni izvor otplate kredita. Kreditni referent određuje koliki će dio prihoda ostati na raspolaganju nakon što se odbiju svi troškovi i uspoređuje to s periodičnim otplatama kredita.

## 4. Uvjeti

Uvjeti se odnose na utjecaj nekih ekonomskih promjena na sposobnost komitenta da nastavi plaćanje kada primjerice neki izvor prihoda nestane kao rezultat opadanja poslovnih aktivnosti.

## 5. Kolateral

Važnost kolaterala je u osiguranju sekundarnog izvora plaćanja. Kolateral može biti imovina financirana traženim kreditom, neka druga imovina koju posjeduje komitent ili osobna garancija supotpisnika kredita.

## 6. Kontrola

Sustavi temeljeni na subjektivnoj procjeni kreditnih referenata obuhvaćaju pravila koja se koriste u ocjeni komitenta koji je aplicirao za kredit. Vrlo često nastaju i razvijaju se u svakoj banci za potrebe banke i pod utjecajem su kreditne kulture svake banke.

Tablica 2. predočuje koji se kriteriji upotrebljavaju u jednoj banci u Japanu, a tablica 3. daje prikaz kriterija jedne financijske institucije u Sjedinjenim Američkim Državama koja odobrava potrošačke kredite za kupovinu automobila.

### Tablica 2: Kriteriji korišteni pri odobravanju potrošačkih kredita jedne banke u Japanu

Maksimalan udio obveza na plaću iznosi 60%
Minimalna starosna dob 25 godina
Dužina vremena na trenutnom poslu min. 2 godine
Tip industrije u kojoj je osoba zaposlena <ul style="list-style-type: none"> <li>• nepoželjno je da je osoba zaposlena u show-business</li> <li>• nepoželjno je da je osoba vozač taksija</li> </ul>

Izvor: Caouette, J.B., Altman, E.I., Narayanan, P., Managing Credit Risk, John Wiley & Sons, New York, 1998, p. 158

### Tablica 3: Kriteriji korišteni pri odobravanju kredita za automobil jedne financijske institucije u Sjedinjenim Američkim Državama

Najmanje jedna godina na postojećoj adresi stanovanja (potreban telefonski broj; potreban dokaz o stambenom statusu)
Trogodišnja povijest stanovanja koja se može provjeriti (za granične slučajeve)
Najmanje jedna godina na postojećem poslu
Trogodišnja povijest zaposlenja koja se može provjeriti (za granične slučajeve)
Zahtjevana dužina zaposlenja
Minimalan mjesečni prihod \$1500 (potrebna je potvrda koja to dokazuje)
Samozaposleni moraju dati potvrdu o prihodu
Udio obveza za plaćanje automobila, hipoteke, osiguranja, tekućih kredita i drugih fiksnih obeza ne smije prelaziti 50%
Čisti rejting kreditnog biroa

Izvor: Caouette, J.B., Altman, E.I., Narayanan, P., Managing Credit Risk, John Wiley & Sons, New York, 1998, p. 158

Drugačiji pristup kod ocjenjivanja kreditne sposobnosti se primjenjuje kada je kredit osiguran. Tako se primjerice kreditori koji nude kreditne linije za kupovinu kuće više oslanjaju na procjenjenu vrijednost kuće i na omjer 'kredit-prema-vrijednosti'<sup>3</sup> koji pokazuje koliko osoba planira pozajmiti u odnosu na vrijednost imovine, a manje na primjerice stabilnost zaposlenja ili tijek gotovine zajmotražitelja.

<sup>3</sup> engleska riječ: loan-to-value

### 4.3. Kredit scoring modeli za stanovništvo

Donošenje odluka o odobravanju potrošačkih kredita može biti podržanom upotrebom subjektivne ocjene kreditnih referenata ili kredit scoring modelima. Budući da potrošački krediti angažiraju manje novca nego li poslovni krediti, ali više kreditnih referenata, sve su više financijske institucije zainteresirane za primjenu kreditnog scoringa koji omogućuje kvalitetnije donošenje odluka i upravljanje kreditnim rizicima uz angažiranje manje ljudi.

Prilikom ocjenjivanja potrošačkih kredita banka može primjenjivati sistem subjektivne procjene kreditnih referenata ili pak kreditni scoring sistem. U oba slučaja kreditni referent prikuplja podatke koji se odnose na karakter, kapacitet i kolateral. Upotrebom isključivo subjektivne analize, kreditni referent subjektivno interpretira informacije slijedeći formalne upute, te na temelju toga prihvaća ili odbija kreditni zahtjev. Ukoliko banka upotrebljava isključivo kvantitativnu analizu, odnosno kreditni scoring, kreditni referent ubacuje tražene podatke u model rezultat kojega je skor odnosno ocjena na temelju koje se donosi odluka o odbacivanju odnosno odobravanju kredita. Najčešće metode koje se upotrebljavaju u kreditnom scoringu su logistička regresija i diskriminacijska analiza. Ostale metode su: probit analiza, matematičko programiranje, neuralne mreže, metoda stabla odlučivanja, Markovljevi lanci, ekspertni sistemi, genetički algoritmi itd. Kredit scoring sistemi kreirani za evaluaciju novih zajmotražitelja su bazirani na dostupnim statističkim podacima koji se odnose na ponašanje komitenata kojima je već ranije odobren kredit. Kako je došlo do porasta potrošačkih kredita banke više nisu mogle obraditi toliku količinu zahtjeva ručno, pa su sve više okrenule kreiranju i upotrebi kredit scoring modela.

Primjer jedne kredit scoring kartice koja je razvijena u University National Bank, a primjenjuje se za kupovinu automobila nalazi se u tablici 4. U pretkoloni se nalaze varijable, a u brojčanom dijelu tablice se nalaze vrijednosti varijabli s pripadajućim bodovima čijim se zbrajanjem dobije kreditni skor za svakog komitenta. Od navedenih 9 varijabli, 5 od njih imaju najveći utjecaj na otplatu potrošakog kredita za automobil, a to su: radni status, stambeno stanje, mjesečni dugovi u odnosu na mjesečni prihod, ukupni prihod i bankovni računi.

**Tablica 4: Primjer kredit skoring kartice korištene za odobravanje kredita za kupovinu automobila**

Godišnji bruto prihod	<10000\$	10000\$ do 20000\$	20000\$ do 40000\$	40000\$ do 60000\$	>60000\$	
	5	15	30	45	60	
Mj.obveze/mj. čisti prihod	> 40%	30-40%	20-30%	10-20%	<10%	
	0	5	20	35	50	
Štednja/čekovni račun	ništa	samo ček. račun	samo štednja	oboje	bez odgovora	
	0	30	30	50	0	
Glavne kreditne kartice	nema	1 ili više	bez odgovora			
	0	30	0			
Kreditna povijest	štetne info zadnjih	unutar 7 godina	nema	info	ispunjene	sve obveze
		-10	0		30	
Dob	≤50 g.	> 50 g.	bez odg.			
	5	25	0			
Stambeni status	najam	posjeduje/kupuje	potpuno vlasništvo	bez odgovora		
	15	40	50	15		
Stabilnost stanovanja	< 1 g.	1-2 g.	2-4 g.	> 4 g.	bez.odg.	
	0	15	35	50	0	
Stabilnost zaposlenja	< 1 g.	1-2 g.	2-4 g.	> 4 g.	nezaposl.	umirovljen
	5	20	50	70	5	70

Izvor: Koch, T.W., MacDonald, S.S., Bank Management, Harcourt College Publishers, Fortworth, Tokyo, 2000, p. 724

Većina banka pri odobravanju potrošačkih kredita koristi kredit skoring modele. Tipično, takvi modeli sadrže varijable prikazane u tablici 5. U tablici su istaknuti atributi svake karakteristike koji imaju pozitivan utjecaj na skor što znači da povećavaju njegovu vrijednost, a samim tim i vjerojatnost za dobivanjem kredita. Tako se primjerice, šanse zajmotražitelja za odobravanjem povećavaju ako posjeduje kuću, ima visok prihod po članu domaćinstva, duže vremena živi na istoj adresi, u braku je, ima malo upita za njegov izvještaj iz kreditnog biroa, posjeduje kreditne kartice, ima čistu kreditnu povijest, posjeduje telefonski broj i/ili je dugo na postojećem poslu.



**Tablica 5: Varijable koje se upotrebljavaju u kredit scoring modelima za odobravanje potrošačkih kredita i njihov utjecaj na skor**

Varijable	Kategorija s pozitivnim utjecajem na skor
Posjeduje li kuću/stan ili ima najam	Posjedovanje
Godine provedene na postojećoj adresi	Što više
Prihod po članu domaćinstva	Visok
Bračni status (samac, u braku, razveden/a, razdvojen/a)	U braku
Zanimanje	Varira
Kredit biro upiti <sup>4</sup>	Malo
Ostale kreditne kartice	Ima
Kreditne kartice naftinih kompanija	Ima
Broj nepovoljnih značajki u kreditnoj povijesti	Malo
Broj ozbiljnih štetnih stavki	Malo
Broj kredit biro upita u posljednjih 6 mjeseci	Malo
Posjeduje li telefonski broj	Da
Broj godina na postojećem poslu	Velik

Izvor: Caouette, J.B., Altman, E.I., Narayanan, P., *Managing Credit Risk*, John Wiley & Sons, New York, 1998, p. 159

Posljednja varijabla ipak postaje sve manje značajna. Naime, kreditori su došli do rezultata da je važnija stabilnost prihoda nego li stabilnost zaposlenja.

Institucija koja želi uvesti kredit scoring sistem izabire između kupovanja generičkog kredit scoring modela od poduzeća koja proizvode i prodaju takve modele ili pak razvija statistički model na temelju svojih podataka. Ako odabere generički proizvod, može se dogoditi da demografska struktura populacije upotrebljena za razvoj takvog modela ne odgovara demografskoj strukturi zajmotražitelja konkretne financijske institucije. Rezultat toga je da performance sistema mogu biti nezadovoljavajuće. Ono što se u praksi događa jest da institucija u prvoj fazi upotrebe kredit scoringa koristi generički sistem, a kasnije kada prikupi dovoljno podataka ide na izgradnju vlastitog scoring modela.

Desai, Crook i Overstreet<sup>5</sup> su istraživali upotrebu diskriminacijske analize i logističke regresije te neuralnih mreža «širenje unatrag» i modularnih neuralnih mreža u

<sup>4</sup> broj potencijalnih kreditora koji su tražili kopiju izvještaja kreditnog biroa za komitenta

<sup>5</sup> Desai, V.S., Crook, J.N., Overstreet, G.A., A Comparison of Neural Network and Linear Scoring Models in Credit Union Environment, *European Journal of Operational Research* 95 (1996), p. 24-35.

izgradnji kredit scoring modela za podatke 3 kreditne zadruge. Također su ispitali performance generičkih modela i usporedili ih s korisniku prilagođenim modelima. Svrha njihovog istraživanja je upotrijebiti podatke iz tri kreditne zadruge kako bi se ispitalo može li se prediktivna moć varijabli povećati ako se regresijska i diskriminacijska analiza zamjene neuralnim mrežama.

Podaci su prikupljeni iz tri kreditne zadruge u SAD-u u periodu između 1988. i 1991. godine. Prva kreditna zadruga sastavljena je od učitelja, druga od zaposlenika telefonskih kompanija, a treća od različitih državnih službenika. Broj kredita u uzorku za prvu zadrugu bio je 505 s 81.58% dobrih kredita, za drugu 762 s 74.02% dobrih i za treću 695 s postotkom dobrih od 78.85%. Modelima je obuhvaćeno 18 prediktorskih varijabli koje su prikazane u tablici 6. Upotrijebljene su varijable koje se inače koriste u kredit scoring modelima za potrošačke kredite, a k tome su još dodane i varijable dobivene iz detaljnih izvještaja kreditnih biroa. Zavisna binarna varijabla označavala je dobre i loše kredite. Kredit je bio svrstan u kategoriju loših, ako je u bilo kojem vremenu u proteklih 48 mjeseci komitentov posljednji kredit bio otpisan ili je komitent bankrotirao. Svi drugi slučajevi su svrstani u kategoriju dobrih.

**Tablica 6: Lista prediktorskih varijabli za odobravanje potrošačkih kredita trima kreditnim zadrugama u SAD-u**

Broj važnijih kreditnih kartica
Posjedovanje kuće
Plaća i ostali prihodi
Postojanje štetnih informacijama i rejting u izvještaju kreditnog biroa
Broj godina provedenih na postojećem poslu
Broj članova obitelji
Broj upita <sup>6</sup> u proteklih 7 mjeseci
Broj mjeseci od kada je komitent otvorio račun
Mjesečna plaćanja (obveze) kao proporcija u mjesečnom prihodu
Računi u kašnjenju u proteklih 12 mjeseci
Ukupni dug kao proporcija prihoda
Dob zajmotražitelja
Broj godina provedenih na postojećoj adresi stanovanja
Broj otvorenih računa na izvještaju kreditnog biroa
Broj aktivnih računa na izvještaju kreditnog biroa
Broj prijašnjih kredita preko kreditne zadruge

Izvor: Desai, V.S., Crook, J.N., Overstreet, G.A., A comparison of neural network and linear scoring models in credit union environment, European Journal of Operational Research 95 (1996), p. 24-35

<sup>6</sup> koliko puta su potencijalni kreditori tražili kopiju izvještaja kreditnog biroa za komitenta

Za korisniku prilagođene modele, linearna diskriminacijska analiza je ispravno klasificirala 82.35% kredita, logistička regresija 82.67%, a neuralne mreže «širenje unatrag» 83.19% kredita. Testiranje značajnosti razlike je pokazalo da razlika nije statistički značajna uspoređujući lineranu diskriminacijsku analizu i logističku regresiju, ali testiranje pokazuje da su neuralne mreže «širenje unatrag» značajno bolje u identificiranju kredita od linearne diskriminacijske analize dok ta razlika u odnosu na logističku regresiju nije značajna. Nadalje, neuralne mreže «širenje unatrag» su uspješno identificirale 49.72% loših kredita., diskriminacijska analiza 38.49% i logistička regresija 44.93%. Razlike su statistički značajne. Kod generičkih modela, linearna diskriminacijska analiza ispravno predviđa 81.12% kredita, a logistička regresija 81.7%, neuralne mreže «širenje unatrag» 80.75%, a modularne neuralne mreže 80.46%. U odnosu na tradicionalne statističke metode ne može se tvrditi da su neuralne mreže «širenje unatrag» signifikantno bolje. Kada se uspoređuju rezultati najboljih neuralnih mreža i najbolje konvencionalne metode, razlika je gotovo zanemariva 0.52%, kada se radi o klasificiranju i dobrih i loših zajedno, no, ta je razlika malo veća 4.79% i na strani neuralnih mreža kada se radi o identificiranju loših kredita. Niti jedna metoda, bilo korisniku prilagođena ili generička, ne identificira loše kredite jednako dobro kao što identificira dobre kredite. Budući da su troškovi davanja kredita lošem komitentu veći nego li odbijanje dobrog komitenta, postotak ispravne klasifikacije loših kredita može biti dobra mjera uspješnosti modela. Autori su u ovom istraživanju pokazali da je logistička regresijska analiza bolja od linearne diskriminacijske analize, i za korisniku prilagođene i za generičke modele. Kada se uspoređuju rezultati korisniku prilagođenih modela s generičkim modelima, pokazuje se da ovi prvi nešto bolje identificiraju loše kredite (razlika je 7.64%) dok je razlika kod identificiranja svih kredita zajedno 2.44% na strani korisniku prilagođenih. Nadalje, korisniku prilagođene neuralne mreže upućuju na obećavajuće područje ako je mjera uspješnosti postotak loših kredita ispravno klasificiranih. Ako je pak mjera uspješnosti postotak ispravno klasificiranih dobrih i loših kredita, logistički regresijski modeli su usporedivi s neuralnim mrežama.

Desai et al.<sup>7</sup> 1997. godine proširuju istraživanje započeto 1996.<sup>8</sup>, primjenjujući genetičke algoritme na iste uzorke o trima kreditnim zadrugama u SAD-u. Glavni cilj istraživanja bio je otkriti može li prediktivna moć varijabli biti povećana ako se statističke metode regresijske i diskriminacijske analize zamjene kombinacijom neuralnih mreža «širenje unatrag» i genetičkim algoritmima za diskriminacijsku analizu. Rezultati pokazuju da je najteže bilo klasificirati slabe kredite. Primjerice, za jednu od kreditnih zadruga, logistička regresija je pogrešno klasificirala u prosjeku 57.16% slabih kredita, zatim 30.76% loših kredita te 14.49% dobrih kredita. Gledajući prosječne rezultate za sve tri kreditne zadruge za 10 uzoraka, logistička regresija je identificira 67.3% kredita ispravno, što je više nego li bilo koji drugi model. Naime, prosječna ispravna klasifikacija diskriminacijske analize iznosi 66.53%, neuralnih mreža «širenje unatrag» 66.38%, a genetičkih algoritama 65.7%. Usporedba pokazuje da je logistička regresija jasno bolja od linearne diskriminacijske analize i genetičkih algoritama, a razlika također postoji i kod usporedbe s neuralnim mrežama «širenje unatrag». Neuralne mreže su se pokazale najboljima kada se radi o ispravnom identificiranju slabih kredita, a genetički algoritmi su najbolji u odnosu na sve ostale kada se radi o korektnoj identifikaciji loših kredita. Genetički algoritmi su u prosjeku ispravno klasificirali 68.38% loših kredita, linearna diskriminacijska analiza 50.82%, logistička regresija 54.32%, a neuralne mreže «širenje unatrag» 50.17%.

Yobas, Crook i Ross<sup>9</sup> uspoređuju prediktivnu sposobnost linearne diskriminacijske analize, neuralnih mreža «širenje unatrag», genetičkih algoritama i stabla odlučivanja u razlikovanju dobrih i sporih komitenata u plaćanju računa kreditnih kartica. Njihovo je istraživanje obuhvaćalo skup podataka od 1001 osobe koja je dobila kreditnu karticu i za koju je poznata kreditna povijest upotrebe te kreditne kartice. Osoba se definira kao spora ako je kasnila u plaćanju jedne ili više otplate u istom periodu, a svi ostali su definirani kao dobri. Upotrebljene varijable se nalaze u tablici 7.

---

<sup>7</sup> Desai, V.S., Conway, D.G., Crook, J.N., Overstreet, G.A., Credit Scoring Models in Credit Union Environment Using Neural Network and Generic Algorithms, IMA Journal of Mathematics Applied in Business & Industry, vol.8, 1997, p.323-346.

<sup>8</sup> Desai, V.S., Crook, J.N., Overstreet, G.A., A Comparison of Neural Network and Linear Scoring Models in Credit Union Environment, op.cit.

<sup>9</sup> Yobas, M.B., Crook, J.N., Ross, P., Credit scoring using neural networks and evolutionary techniques, IMA Journal of Mathematics Applied in Business and Industry, vol.11, 2000, p.111-125.

**Tablica 7: Upotrebljene varijable kod odobravanja potrošačkih kredita putem kreditne kartice**

Status zaposlenja komitenta
Broj godina u banci
Vrijednost hipoteke na kuću
Broj djece
Broj godina provedenih u poduzeću gdje je trenutno zaposlen
Stambeni status
Tipovi računa otvorenih u banci
Ostale kartice koje posjeduje
Troškovi
Procjenjena vrijednost kuće
Telefon u kući
Prihod komitenta
Prihod bračnog partnera
Glavne kreditne kartice koje posjeduje

Izvor: Yobas, M.B., Crook, J.N., Ross, P., Credit scoring using neural networks and evolutionary techniques, IMA Journal of Mathematics Applied in Business and Industry, vol.11, 2000, p.111-125

Linearna diskriminacijska analiza je uspješno klasificirala 68.4% svih kredita<sup>10</sup> (preuzeto iz istraživanja provedeno 1992.). Neuralne mreže «širenje unatrag» su ispravno klasificirale 64.2% kredita što je najveći dobiveni postotak u ovom istraživanju. Postotak ispravne klasifikacije za stablo odlučivanja iznosi 62.3% što je dobiveno kao srednja proporcija za 10 stabala odlučivanja. Genetički algoritmi pravilno raspoređuju 64.5% što je najveća srednja vrijednost za 50 pokušaja u svakom eksperimentu.

Galindo i Tamayo<sup>11</sup> na skupu podataka o hipotekarnim kreditima ispituju prednosti i nedostatke logističke regresije, neuralnih mreža «širenje unatrag», metoda stabla odlučivanja i k-najbližih susjeda (k-nearest neighbours) u procjenjivanju rizika hipotekarnih kredita za kuće. Autori ističu da je potrebno provesti puno ispitivanja različitih kombinacija algoritama u cilju pronalaženja najboljeg rješenja. Oni stoga, za kompleksne realne podatke, predlažu interdisciplinarni pristup koji kombinira statistiku i algoritme za strojno učenje. Postoje mnogo algoritama koji su dostupni za izgradnju modela tako da je zapravo glavni problem u praksi, odabiranje i kombiniranje različitih algoritama. Podaci na kojima je provedeno

<sup>10</sup> izračunato uporebom 'jackknife' metode

<sup>11</sup> Galindo, J., Tamayo, P., Credit risk assessment using statistical and machine learning: Basic methodology and risk modeling applications, Computational Economics, 15, 2000, p. 107-143

ovo istraživanje su sadržavali 4000 hipotekranih kredita jedne banke. Za svaki kredit na raspolaganju su bile 24 varijable koje se nalaze u tablici 8. Zavisna varijabla binarnog je karaktera gdje jedna kategorija predstavlja kredite u kašnjenju, a to su oni kod kojih nije bilo plaćanja u posljednja 2 mjeseca. Ostali krediti su svrstani u grupu u kojoj nije bilo kašnjenja.

Varijable koje su činile probit model su bile slijedeće:

- Pokriva li vrijednost garancije barem 100% ukupnog duga
- Pripada li kredit u posebnu kategoriju s nižim kamatama zbog lošijeg ekonomskog statusa domaćinstva
- Je li je kredit namjenjen za izgradnju nove kuće
- Indeks

Test greška za probit model iznosi 15.13%. Kod usporedbe stvarnog stanja i rezultata predviđanja dobivenih probit modelom zabilježena je ukupna greška od 15.8%. Pogrešna klasifikacija za kredite koji nisu kasnili u plaćanju iznosi 6.10%, a koji su kasnili iznosi 25.2%.

Rezultati stabla odlučivanja izdvajaju slijedeće varijable koje su se pokazale najutjecajnijima za otplatu kredita:

- Ukupan dug
- Plaćanje rate u prvom mjesecu
- Pokriva li vrijednost garancije barem 100% ukupnog duga
- Iznos neplaćene razlika koja je trebala biti plaćena
- Indeks

Optimalna veličina stabla sadržavala je 120 grana s prosječnom test greškom od 8.3%. Ukupan postotak pogrešne klasifikacije iznosi 9.10%. Stopa pogrešne klasifikacije za kredite kod kojih nije bilo kašnjenja iznosi 11.99%, a za kredite kod kojih je bilo kašnjenja 6.29%. Najbolje rezultate daje stablo odlučivanja sa 120 grana koje sadžava prosječnu test grešku od 8.3% na uzorku od 2000 podataka. Nakon toga slijede neuralne mreže sa 16 skrivenih slojeva trenirane u 80 iteracija. Prosječna test greška iznosila je 11% na 2000 podataka. Prosječna test greška za metodu k-najbližih susjeda iznosila je 14.95%, a za probit model 15.13%. Autori predlažu kombiniranje dobivenih modela kako bi se postigla veća preciznost upotrebom logičkih operatora AND i OR. Tako su oni pokazali da je za njihove podatke za predviđanje kredita koji neće kasniti najbolja upotreba kombinacije stabla odlučivanja

i probit modela, a za predviđanje kredita koji će kasniti upotreba stabla odlučivanja ili neuralnih mreža.

**Tablica 8: Varijable upotrebljene za analiziranje kašenjenja u otplati hipotekarnih kredita za kuće**

Veličina kredita
Iznos još neplaćene razlike
Iznos neplaćene razlike koja je trebala biti plaćena
Ukupni dug (suma prethodne dvije varijable)
Vrijednost garancije odnosno hipoteke
Pokriva li vrijednost garancije barem 100% ukupnog duga
Pokriva li vrijednost garancije barem 200% ukupnog duga
Pripada li kredit u posebnu kategoriju s nižim kamatama zbog lošijeg ekonomskog statusa domaćinstva
Radi li se o regularnom kreditu (bez ikakvih popusta)
Je li je kredit namjenjen za kupovinu već postojeće kuće
Je li je kredit namjenjen za izgradnju nove kuće
Je li je kredit namjenjen za adaptaciju kuće
Je li je kredit namjenjen za kupovinu ili konstukciju kuće
10 varijabli sadrži informacije o kreditnoj povijesti u razdoblju od 10 mjeseci (1 označava da rata nije plaćena za taj period, a 0 da jest). Na temelju toga se kreira matrica koja sarži 0 i 1 za prvih 10 mjeseci svakog računa
Kreditna povijest i vjerojatnost plaćanja izražena u $P_{ij}$ koje se definira kao vjerojatnost da će se stanje računa u danom periodu promijeniti iz stanja $i$ u stanje $j$ : $P_{ij} = P(\text{stanje } t = j \mid \text{stanje } t-1 = i)$ Varijabla Index se kreira na slijedeći način: $Index = \frac{\sum_{k=1}^{10} P_{i1}^k}{10}$ gdje $P_{i1}^k$ zauzima vrijednost $P_{i1}^{11-k}$ ako je račun u stanju $i$ u mjesecu $k$ .

Izvor: Galindo, J., Tamayo, P., Credit risk assessment using statistical and machine learning: Basic methodology and risk modeling applications, Computational Economics, 15, 2000, p. 107-143

Piramuthu<sup>12</sup> se bavi upotrebom neuralnih mreža i fuzzy sistema pri donošenju kreditnih odluka. Naime, teško je objasniti ono što leži iza odluka dobivenih neuralnim mrežama pa stoga istraživači razvijaju metode upotrebom neuralnih mreža kako bi proizveli pravila koja se onda koriste za racionalno objašnjavanje odluka. Zato se razvijaju 'neurofuzzy' sistemi koji imaju karakteristike i neuralnih i 'fuzzy' sistema. Autor je primijenio 'neurofuzzy' sisteme na podatke o odobravanju kreditne kartice za

<sup>12</sup> Piramuthu, S., Financial Credit-risk Evaluation with Neural and Nerofuzzy Systems, European Journal of Operational Research, 112, 1999, p.310-312.

potrošačke kredite. Podaci su sadržavali 9 diskretnih i 6 kontinuiranih atributa. Ukupan broj obuhvaćenih kredita bio je 653. Od toga je 296 odobrenih kredita, a ostalih 357 je odbijeno. Skup podataka je podjeljen na 490 podataka koji su upotrebljeni za treniranje, a od toga je 268 odobrenih i 222 su odbijena. Skup od 163 podataka je upotrebljen za testiranje. Od toga je 89 odobrenih te 74 odbijena kredita. Neuralne mreže u kojima je bilo 15 ulaznih neurona, 8 skrivenih i 1 izlazni su imale stopu učenja 0.19. Mreža je radila sve dok ukupna suma kvadrata greške nije dosegla 0.04 ili dok broj iteracija ne dosegne 2000. Za 'neurofuzzy' sisteme u kojima je bilo 15 ulaznih neurona, 250 skrivenih i 1 izlazni stopa učenja bila je 0.015, a sistem je radio sve dok ukupna suma kvadrata greške nije bila 0.04 ili dok broj iteracija nije bio 2000. Prosječna<sup>13</sup> točnost klasifikacije za neuralne mreže za trening podatke je iznosila 95.59% (0.53), a za test podatke 83.56% (7.22). Za 'neurofuzzy' sisteme prosječna točnost klasifikacije za trening podatke je iznosila 91.74% (1.20), a za test podatke 77.91% (5.10).

Istraživanje koje je napravio West<sup>14</sup> ispituje mogućnost poboljšanja preciznosti kredit scoring modela testiranjem 5 različitih arhitektura neuralnih mreža na dva stvarna skupa podataka. Napravljeno je 10 ponavljanja za svaki pokušaj neuralnih mreža, a rezultati iz vanjskih nezavisnih uzoraka su testirani McNemar hi-kvadrat testom<sup>15</sup> kako bi se vidjelo jesu li razlike statistički značajne. Ti su rezultati uspoređeni s rezultatima dvije parametarske metode: 1.) linearnom diskriminacijskom analizom i 2.) logističkom regresijom i dvije neparametarske metode: 1.) k-najbližih susjeda i 2.) stablom odlučivanja.

Testirane su slijedeće neuralne mreže: 1.) Mreža «širenje unatrag», 2.) Modularna mreža, 3.) Mreža s radijalno zasnovanom funkcijom, 4.) Mreža učeće vektorske kvantizacije, 5.) Mreža 'fuzzy' adaptivne rezonancije.

Jedan skup podataka (njemački podaci) sastojao se od 700 kreditno sposobnih i 300 kreditno nesposobnih komitenata, a drugi skup podataka (australski podaci) se sastojao od 307 prvih i 383 drugih zajmotražitelja. Za svakog komitenta su bile na raspolaganju 24 varijable kojima se opisivala kreditna povijest, stanje računa,

---

<sup>13</sup> neuralne mreže i neurofuzzy sistemi su pokrenuti 10 puta. Izračunata je prosječna vrijednost na temelju svih 10 ponavljanja

<sup>14</sup> West, D., Neural Network Credit Scoring Models, Computers & Operations Research, 27, 2000, p. 1131-1152.

<sup>15</sup> Sheskin, D.J. 1997, Handbook of Parametric and Nonparametric Statistical Procedures, CRC Press, Washington D.C., 1997



namjena kredita, količina kredita, status zaposlenja, osobne informacije, godine, stambeni status i posao. Podaci su se sastojali od 6 kontinuiranih i 8 kategorijalnih varijabli. Niti jedna varijabla nije sadržavala podatke kreditnih biroa.

Za neuralne mreže je broj ulaznih neurona bio broj varijabli u skupu podataka. Izlazni neuron je imao dva stanja koja su predstavljala kreditnu odluku. Inicijalni eksperimenti su imali dužinu treniranja koja se kretala od 30 tisuća do 300 tisuća iteracija. Opterećenja mreže su mijenjana nakon svake iteracije učenja. Parametri treniranja mreža su smanjivani svakih 10 iteracija učenja, a treniranje je prekidano nakon 50 iteracija učenja. Napravljeno je 10 ponavljanja za svaki od 10 podjela kako bi se smanjila stohastička varijabilnost procesa treniranja mreže. Za metodu "k najbližih susjeda", za njemačke podatke  $k=3$ , a za australske  $k=7$ . Veličina stabla odlučivanja je limitirana hi-kvadrat pravilom za  $p=0.05$ . Granična vrijednost upotrebljena za razlikovanje između dvije kreditne grupe za kvantitativne metode je 0.5.

Ako se kod neuralnih mreža u razmatranje uzmu 3 najbolja dobivena rezultata, tada se vidi da su modularne mreže ne samo najbolje među neuralnim mrežama već da su i najpreciznije od svih kvantitativnih metoda upotrebljenih u ovom istraživanju. Za skup podataka iz Njemačke greška modularne mreže iznosi 0.2243 i značajno je manja od greške dobivene s mreža s radijalno zasnovanom funkcijom koja iznosi 0.2437, mreže «širenje unatrag» 0.2496 i logističke regresije koja daje grešku od 0.2370. Za ovaj skup podataka cjelokupna greška za modularnu mrežu bilježi poboljšanje u preciznosti za približno 0.005 u odnosu na linearnu diskriminacijsku analizu i 0.013 u odnosu na logističku regresiju. Za australske podatke najpreciznija je mreža s radijalno zasnovanom funkcijom s greškom od 0.1222 zatim mreža «širenje unatrag» s 0.1232 i MOE s 0.1239. Kako bi se ispitala statistička signifikantnost razlika između kreditnih modela upotrebljen je McNemarov test. Za njemačke kredite test pokazuje da su modularna mreža, mreža s radijalno zasnovanom funkcijom, mreža «širenje unatrag» i logistička regresija dali modele koji se statistički ne razlikuju od najpreciznijeg modela. Za australske kredite ovim spomenutim treba još dodati linearnu diskriminacijsku analizu. Treba napomenuti da su se sve metode pokazale lošijima kod klasificiranja loših kredita odnosno kod identifikacije zajmotražitelja koji nisu kreditno sposobni.

#### 4.4. Reference

1. Caouette, J.B., Altman, E.I., Narayanan, P., *Managing Credit Risk*, John Wiley & Sons, New York, 1998.
2. Desai, V.S., Crook, J.N., Overstreet, G.A., *A Comparison of Neural Network and Linear Scoring Models in Credit Union Environment*, *European Journal of Operational Research* 95 (1996), p. 24-35.
3. Desai, V.S., Conway, D.G., Crook, J.N., Overstreet, G.A., *Credit Scoring Models in Credit Union Environment Using Neural Network and Generic Algorithms*, *IMA Journal of Mathematics Applied in Business & Industry*, vol.8, 1997, p.323-346.
4. Galindo, J., Tamayo, P., *Credit risk assessment using statistical and machine learning: Basic methodology and risk modeling applications*, *Computational Economics*, 15, 2000, p. 107-143
5. Hand, D.J., Henley, W.E., *Statistical Classification Methods in Consumer Credit Scoring: a Review*, *Journal Royal Statistical Society A*, 160, 1997, p. 523-541
6. Koch, T.W., MacDonald, S.S., *Bank Management*, Harcourt College Publishers, Fortworth, Tokyo, 2000.
7. Piramuthu, S., *Financial Credit-risk Evaluation with Neural and Neurofuzzy Systems*, *European Journal of Operational Research*, 112, 1999, p.310-312.
8. Sheskin, D.J. 1997, *Handbook of Parametric and Nonparametric Statistical Procedures*, CRC Press, Washington D.C., 1997
9. West, D., *Neural Network Credit Scoring Models*, *Computers & Operations Research*, 27, 2000, p. 1131-1152.
10. Yobas, M.B., Crook, J.N., Ross, P., *Credit scoring using neural networks and evolutionary techniques*, *IMA Journal of Mathematics Applied in Business and Industry*, vol.11, 2000, p.111-125.