

بررسی میزان کارایی مدل‌های شناسایی الگو در طراحی و ساخت مدل‌های امتیازبندی اعتباری

محمد فریدون کیانی

مهری محفوظیان

رئوس مطالب

- معرفی امتیازبندی اعتباری
- مدل‌های امتیازبندی اعتباری
- روش‌های ارزیابی مدل‌ها
- مشکلات و نقاط ضعف هر کدام
- تابع هزینه
- نتایج عملی
- بحث و نتیجه‌گیری

امتیازبندی اعتباری

- بحرانهای مالی اخیر در سطح بین المللی
 - بدترین بحران بعد از ۱۹۲۰
 - میزان زیان در فوریه ۲۰۰۸ بالغ بر ۱۴۸ میلیارد دلار [1]
- منشاء: عدم توجه کافی به بحث مدیریت ریسک در بانکها و موسسات مالی
 - اعطای حجم بالایی از تسهیلات پرخطر بدون ارزیابی درست میزان توانایی افراد متقاضی
- لزوم توجه هر چه بیشتر به مدیریت ریسک و ابزارهای کنترل و سنجش آن (توصیه نامه بال [2])
 - توجه بیشتر به پیش بینی های شرکتهای رتبه بندی (Standard & Poor, Moody's KMV)
 - الزام به طراحی و بکارگیری یک سیستم رتبه بندی داخلی (IRB)

امتیازبندی اعتباری

- بانکها ماشینهای ریسک پذیر
 - خطرپذیری لازمه سودآوری در آینده
 - مواجهه با انواع ریسکها
 - ریسک اعتباری
 - احتمال قصور در بازپرداخت تعهدات
 - کسانیکه نتوانند تعهدات خود را در طی ۹۰ روز به انجام برسانند [3].
 - روشهای سنتی
 - تجربه و تخصص افراد متخصص (معیار 5C)
 - » شخصیت، سرمایه، وثیقه، ظرفیت و شرایط
 - خطای انسانی، سلیقه ای، کمبود افراد متخصص
 - روشهای نوین (سیستم امتیازبندی اعتباری)
 - پیش بینی خودکار آینده مالی با تکنیک مشتریان در دو گروه خوب و بد
 - قابلیت اعمال روشهای شناسایی الگو
- تخمین درست تابع f که کار نگاشت ورودیها (ویژگیهای مشتری) را به خروجی (کلاس مشتری) به انجام می رساند.

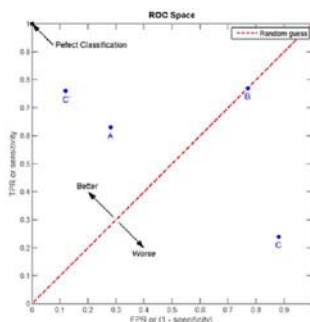
$$f(x_1, x_2, \dots, x_m) = y$$

امتیازبندی اعتباری

- **پیدایش امتیازبندی اعتباری**
 - دوران [4] در ۱۹۴۱ استفاده از تحلیل ممیز برای جداسازی وام گیرنده های خوب و بد
 - اولین سیستم خیره در امتیازبندی اعتباری فیر و آیزاک ۱۹۵۰ [4]
 - ۱۹۶۰ کارتهای اعتباری، معرفی شاخص احتمال قصور (PD)
- **روشهای امتیازبندی اعتباری**
 - آماری (تحلیل ممیز، لاجیت و پروبیت)
 - داده کاوی و هوشمند (شبکه های عصبی، درخت تصمیم، ماشین بردار پشتیبان)
 - مبتنی بر منطق فازی
 - ترکیبی

روشهای ارزیابی میزان کارایی

		Actual Class	
		p	n
Predicted Class	p	True Positives	False Positives
	n	False Negatives	True Negatives
Column Totals:		P	N



- **روشهای ارزیابی میزان کارایی**
 - خطای آزمایشی (Error)
 - **Confusion Matrix**
 - خطای نوع اول (FN)
 - تعداد نمونه های منفی (مشتریان خوب) که به اشتباه دسته بندی شده اند.
 - خطای نوع دوم (FP)
 - تعداد نمونه های مثبت (مشتریان بد) که به اشتباه دسته بندی شده اند.
 - نمودار ROC و سطح زیر منحنی (AUC)
 - توجه یکسان به خطای نوع اول و نوع دوم
 - کارایی پایین در مدلهایی که بصورت گسسته عمل می کنند.

تابع هزینه

- در امتیازبندی اعتباری هزینه دسته بندی نادرست هر کدام از نمونه ها یکسان نیست.
 - هر دو خطای نوع اول و دوم باید مد نظر قرار داده شود.
 - میزان هزینه هر کدام مشخص گردد.
 - نسبت تعداد نمونه های هر دسته در نظر گرفته شود.
- (D. West, 2000) هزینه دسته بندی نادرست در امتیازبندی اعتباری را ۵ به ۱ تعیین کرد.

$$\text{Cost}_{\text{Function}} = C_{12} \pi_2 \text{FP} + C_{21} \pi_1 \text{FN}$$
 - $C_{12} = 5, C_{21} = 1$
 - π_1 نسبت نمونه های منفی (مشتریان خوب) به کل نمونه ها
 - π_2 نسبت نمونه های مثبت (مشتریان بد) به کل نمونه ها

پیاده سازی مدلها

- روشهای بکاررفته
 - آماری
 - لاجیت (Logit)
 - پروبیت (Probit)
 - داده کاوی و هوشمند
 - نزدیکترین همسایه (kNN)
 - بیزین ساده (Naïve Bayes)
 - درخت تصمیم (Decision tree)
 - شبکه های عصبی پیشران (Feed Forward Neural Networks)
 - ماشین بردار پشتیبان (Support Vector Machine)
- معیارهای ارزیابی
 - خطای تست (Er)
 - خطای نوع اول (FN)
 - خطای نوع دوم (FP)
 - سطح زیرمنحنی (AUC)
 - تابع هزینه (Cost_Fn)

مجموعه داده ها

— مجموعه داده ها

- داده های محک از مجموعه داده های دانشگاه کالیفرنیا (UCI)
 - مجموعه داده های اعتباری استرالیا و آلمان
- مجموعه داده های اعتباری مشتریان حقوقی بانک کارآفرین در فاصله ۱۰ سال
 - مجموعه داده ایران

مجموعه داده	تعداد رکوردها	تعداد ویژگیها	نسبت تعداد مشتریان بد	نوع داده ها	فیلد خالی
آلمان	۱۰۰۰	۲۰	%۳۰	حقیقی	ندارد
استرالیا	۶۹۰	۱۴	%۵۰	حقیقی	دارد
ایران	۱۹۹۴	۶۶	%۱۵	حقوقی	دارد

Logit & Probit

	Fun	Error	FN	FP	AUC	Cost_Fn
German	Logit	0.2393	0.1124	0.5356	0.7813	0.882
	Probit	0.2387	0.1114	0.5356	0.7824	0.8813
Australia	Logit	0.1408	0.1298	0.1543	0.9198	0.4155
	Probit	0.1388	0.1333	0.1457	0.9202	0.398
Iran	Logit	0.1544	0.0127	0.96	0.6046	0.7308
	Probit	0.154	0.0114	0.965	0.6053	0.7335

K-Nearest Neighbor

	k	Error	FN	FP	AUC	Cost_Fn
German	1	0.282	0.1848	0.5089	0.6532	0.8927
	3	0.258	0.1286	0.56	0.6557	0.93
Australia	1	0.3476	0.2825	0.4283	0.6446	1.11
	3	0.3214	0.2211	0.4457	0.6666	1.114
Iran	1	0.2303	0.1402	0.7425	0.5586	0.676
	3	0.2232	0.1336	0.7325	0.5669	0.663

Naïve Bayesian

		Error	FN	FP	AUC	Cost_Fn
German	Uniform	0.3127	0.3371	0.2556	0.7037	0.6193
	Empirical	0.2757	0.2271	0.3889	0.692	0.7423
Australia	Uniform	0.1772	0.0842	0.2924	0.8117	0.6973
	Empirical	0.1806	0.0798	0.3054	0.8074	0.7239
Iran	Uniform	0.3133	0.2138	0.8788	0.5244	0.8408
	Empirical	0.3099	0.2062	0.9	0.5165	0.8502

Feed-Forward Neural Network

	Error	FN	FP	AUC	Cost_Fn
German	0.258	0.1219	0.5756	0.6513	0.9487
Australia	0.1476	0.1325	0.1663	0.8506	0.4435
Iran	0.1553	0.009	0.9875	0.5365	0.7483

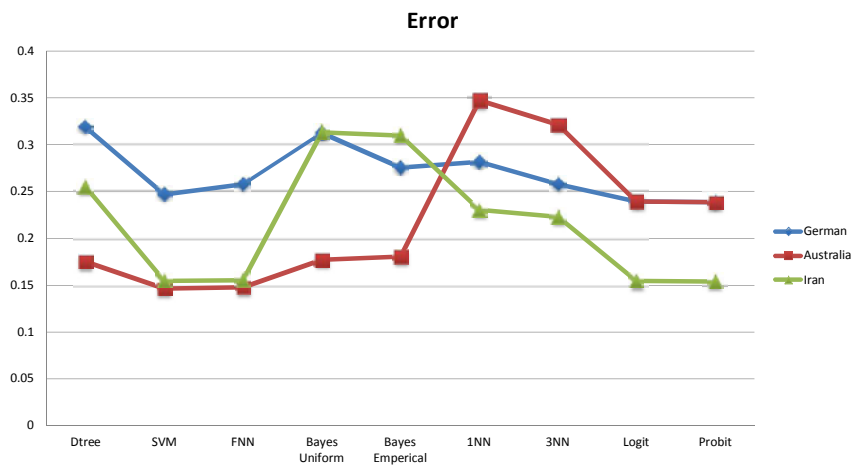
Support Vector Machine (SVM)

	Error	FN	FP	AUC	Cost_Fn
German	0.247	0.1143	0.5567	0.6645	0.915
Australia	0.1466	0.1991	0.0815	0.8597	0.2919
Iran	0.1546	0.0062	0.9988	0.4975	0.7542

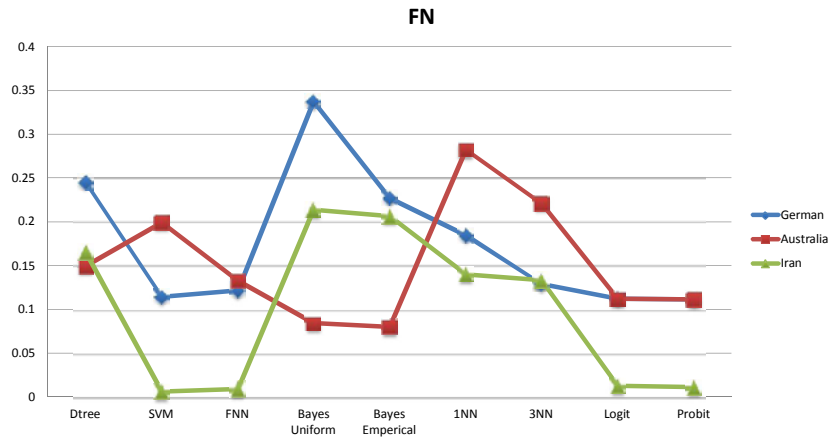
Decision Tree

	Error	FN	FP	AUC	Cost_Fn
German	0.3190	0.2448	0.4922	0.6315	0.9097
Australia	0.1752	0.1491	0.2076	0.8216	0.5447
Iran	0.2548	0.1657	0.7613	0.5365	0.7118

خطای آزمایشی



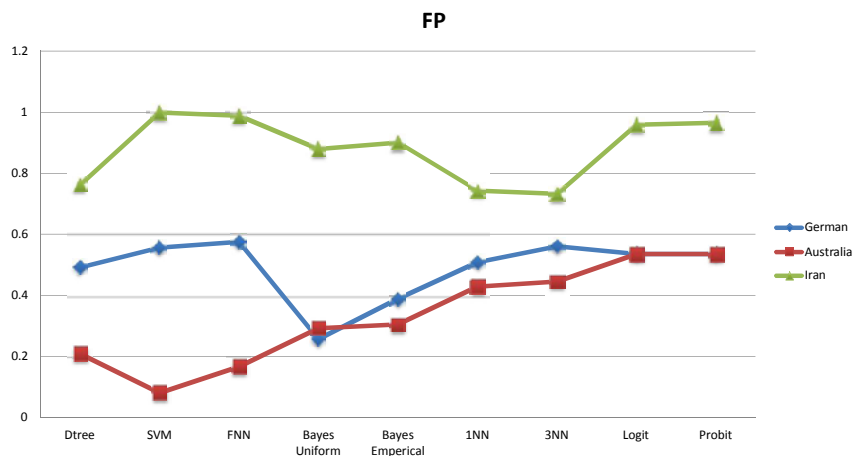
خطای نوع اول



اولین همایش بین المللی بانکداری الکترونیک و نظامهای پرداخت

17

خطای نوع دوم

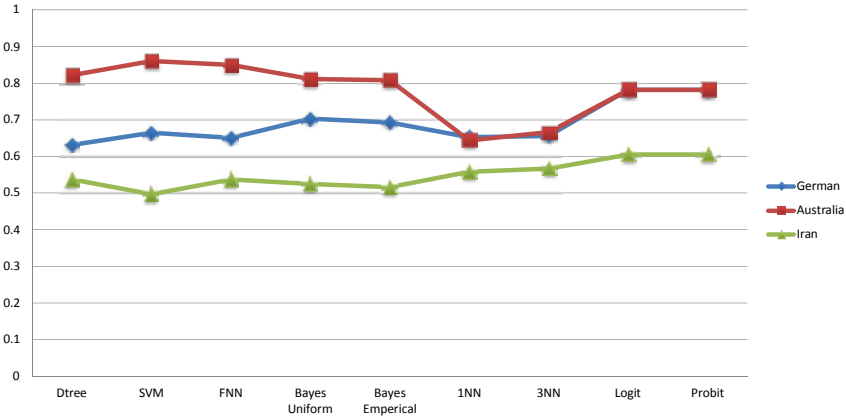


اولین همایش بین المللی بانکداری الکترونیک و نظامهای پرداخت

18

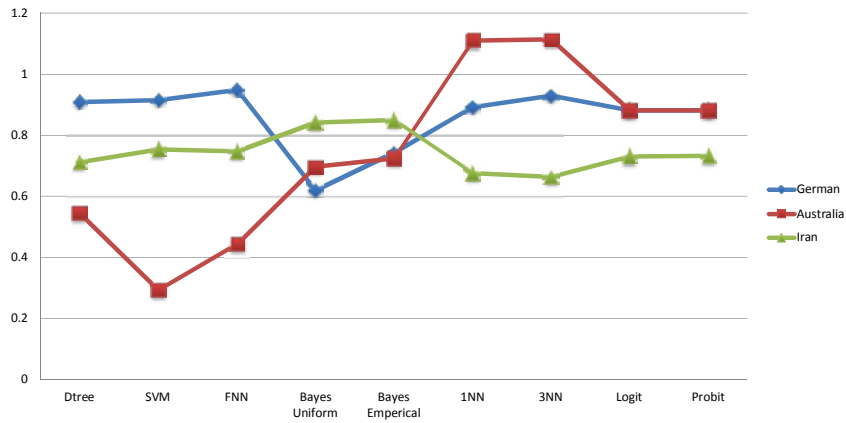
سطح زیر منحنی

AUC

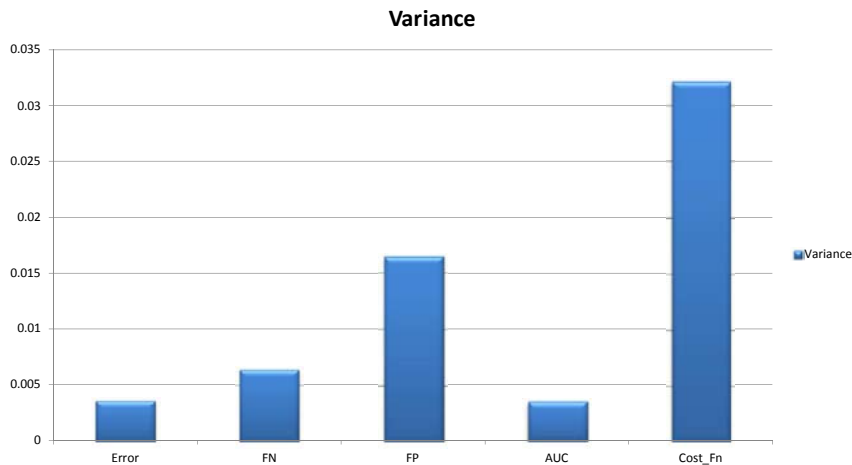


تابع هزینه

CostFn



میانگین واریانس



اولین همایش بین المللی بانکداری الکترونیک و نظامهای پرداخت

21

بحث و نتیجه گیری

- در این مطالعه
 - هفت روش برای ایجاد مدل‌های امتیازبندی اعتباری
 - سه مجموعه داده اعتباری مختلف
 - پنج معیار ارزیابی
- از نظر کیفی معیار خطای آزمایشی ضعیفترین نتایج را بخصوص در داده‌هایی که دچار عدم توازن در تعداد نمونه‌های هر دسته هستند دارد.
- خطای نوع اول و نوع دوم بخاطر توجه یک بعدی به مسئله نمی‌توانند به تنهایی معیار مناسبی برای سنجش کارایی مدل‌ها تلقی شوند.
- سطح زیر منحنی علیرغم توجه به هر دو نوع خطای نوع اول و دوم قابلیت تفکیک‌کنندگی کمی را بخاطر توجه یکسان به هر دو نوع خطا دارد که در نتیجه نمی‌تواند معیار مناسبی برای مسائل مبتنی بر تابع هزینه مانند امتیازبندی اعتباری باشد.
- تابع هزینه
 - قابلیت اعمال محدودیتهای چندگانه
 - تعیین میزان اهمیت هر کدام از محدودیتهای دلخواه
 - قابلیت جداسازی بالاتر نسبت به دیگر

اولین همایش بین المللی بانکداری الکترونیک و نظامهای پرداخت

22

بحث و نتیجه گیری

- نحوه عملکرد هر مدل وابسته به نوع داده های بکار رفته در آن است.
 - داده های استرالیا تفاوت چندانی بین خطای نوع اول و نوع دوم وجود ندارد.
 - در حالیکه در داده های آلمان و بخصوص داده های ایران تفاوت کاملاً محسوس و تابعی از نسبت داده های است.
- نمی توان هیچ مدلی را انتخاب نمود که در همه مجموعه داده ها دارای بهترین عملکرد باشد.
- اهمیت انتخاب روش ارزیابی مناسب نسبت به انتخاب خود مدل
- با انتخاب یک تابع هزینه مناسب با اطمینان بیشتری در خصوص میزان کارایی مدل‌های مورد مطالعه اقدام نمود که قدم اول در طراحی هر مدل شناسایی الگو قلمداد می شود.
- کارهای آینده
 - دسته بندی چندکلاسی از روی داده های دو کلاسی
 - قابلیت تفسیرپذیری پایین مدلها

منابع

- A. J. McNeil, J. P. Wendin. (2007). Bayesian inference for generalized linear mixed models of portfolio credit risk. *Journal of Empirical Finance*, 14, 131-149.
- A. Khashman. (2010). Neural networks for credit risk evaluation: Investigation of different neural models and learning schemes. *Expert Systems with Applications*, 37, 6233-6239.
- C. -F. Tsai. (2008). Financial decision support using neural networks and support vector machines. *Expert Systems*, 25 (4), 380-393.
- C. J.C. Burges. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, 2, 121-167.
- D. J. Hand, N. M. Adams. (2000). Defining attributes for scorecard construction in credit scoring. *Journal of Applied Statistics*, 27 (5), 527-540.
- D. West. (2000). Neural network credit scoring models. *Computers & Operations Research*, 27, 1131-1152.
- G. Donga, K. K. Laib, J. Yen. (2010). Credit scorecard based on logistic regression with random coefficients. *Procedia Computer Science*, 1, 2463-2468.
- G. Löffler, P. N. Posch, C. Schone. (2004). *Bayesian methods for improving credit scoring models*. Department of Finance, University of Ulm, Helmholtzstrasse, Ulm, Germany.
- H. Abdou, J. Pointon, A. El-Masry. (2008). Neural nets versus conventional techniques in credit scoring in Egyptian banking. *Expert Systems with Applications*, 35, 1275-1292.
- H. -C. Yeh, M. -L. Yang, L. -C. Lee. (2007). An empirical study of credit scoring model for credit card. *IEEE*.
- J. Kim. (2005). *A Credit Risk Model for Agricultural Loan Portfolios Under the New Basel Capital Accord*. PhD thesis, Texas A&M University.
- K. Komorad. (2002). *On Credit Scoring Estimation*. Master's Thesis presented to obtain the Master of Science degree, Institute for Statistics and Econometrics Humboldt University, Berlin.

منابع

- M. Daubie, P. Levecq, N. Meskens. (2002). A comparison of the rough sets and recursive partitioning induction approaches: An application to commercial loans. *International Transactions in Operational Research*, 9, 681-694.
- M. F. Kiani, F. Mahmoudi. (2010). A new hybrid method for credit scoring based on clustering and support vector machine (ClSVM). *International Conference on Financial Engineering (ICF, 2010)*, (pp. 585-589).
- M.-H. Jiang, X.-C. Yuan. (2007). Construction and Application of PSO-SVM Model for Personal Credit Scoring. *Springer-Verlag Berlin Heidelberg, ICCS 2007, Part IV, LNCS 4490*, (pp. 158-161).
- P. Giudici. (2001). Bayesian data mining, with application to benchmarking and credit scoring. *Applied Stochastic Models in Bussiness and Industry*, 17, 69-81.
- P.-H. Chen, C.-J. Lin, B. Scholkopf. (2005). A tutorial on n-support vector machines. *APPLIED STOCHASTIC MODELS IN BUSINESS AND INDUSTRY*, 21, 111-136.
- S. J. Huang, J. Yu. (2010). Bayesian analysis of structural credit risk models with micro structure noises. *Journal of Economic Dynamics & Control*, 34, 2259-2272.
- T. Mahlmann. (2004). *Classification and rating of firms in the presence of financial and non-financial information*. Chair of Banking, University of Cologne.
- T.-S. Leea, C.-C. Chiub, Y.-C. Chouc, C.-J. Lud. (2006). Mining the customer credit using classification and regression tree and multivariate adaptive regression splines. *Computational Statistics & Data Analysis*, 50, 1113-1130.
- V. L. Plantamura, B. Soucek, G. Visaggio. (1993). The holographic fuzzy learning for credit scoring. *Proceedings of 1993 International Joint Conference on Neural Networks*, (pp. 729-732).
- W. Abramowicz, M. Nowak and J. Szykiel. (2003). *Bayesian Networks as a Decision Support Tool in Credit Scoring Domain*. The Poznan University of Economics.
- W. Greene. (1998). Sample selection in credit-scoring models. *Japan and the World Economy*, 10, 299-316.
- X. Wu. (2008). *Credit Scoring Model Validation*. Master Thesis, Korteweg-de Vries Institute for Mathematics.
- X. Zhou, D. Zhang, Y. Jiang. (2008). A New Credit Scoring Method Based on Rough Sets and Decision Tree. *Springer-Verlag Berlin Heidelberg*, (pp. 1081-1089).