Profit-Oriented Feature Selection in Credit Scoring Applications

Nikita Kozodoi[1,2], Stefan Lessmann[1], Konstantinos Papakonstantinou[2], and Bart Baesens[3]

[1] Humboldt University of Berlin, Berlin, Germany
[2] Kreditech Holding, Hamburg, Germany
[3] Catholic University of Leuven, Leuven, Belgium

nikita.kozodoi@hu-berlin.de

Presentation Structure

1. Background
2. Research Motivation
3. Related Literature
4. Empirical Experiments
5. Conclusions

Background

Credit scoring:

- the use of **statistical models** to support decision-making in the **retail credit sector** *(Crook et al. 2007)*
- classification task of distinguishing **BAD** and **GOOD** loans
- **scorecard** - model that estimates probability of default

Background

Credit scoring:

- the use of **statistical models** to support decision-making in the **retail credit sector** *(Crook et al. 2007)*
- classification task of distinguishing **BAD** and **GOOD** loans
- **scorecard** - model that estimates probability of default

Feature Selection:

- removing **redundant** and **irrelevant** features to improve the performance of the scorecard and its interpretability
- helps **reducing costs** of gathering and storing customer data
- facilitates **comprehensible models** enforced by regulation

Research Motivation

Standard FS approaches:

- **filters**, **wrappers** and **embedded methods** *(Guyon et al. 2006)*
- using **statistical criteria** to select features:
    - correlation
    - Fisher score
    - information gain
    - area under the ROC curve
    - predictive accuracy
    - etc.

filters

wrappers

Research Motivation

Credit scoring literature:

- using **profit-driven indicators** instead of standard measures improves scorecard profitability *(Finlay 2010, Verbraken et al. 2014)*
- research on profit-driven FS in credit scoring is limited to the **embedded SVM framework** *(Maldonado et al. 2015, 2017)*
- however, **SVM perform poorly** compared to other algorithms in credit scoring task *(Lessmann et al. 2015)*

Research Motivation

Credit scoring literature:

- using **profit-driven indicators** instead of standard measures improves scorecard profitability *(Finlay 2010, Verbraken et al. 2014)*
- research on profit-driven FS in credit scoring is limited to the **embedded SVM framework** *(Maldonado et al. 2015, 2017)*
- however, **SVM perform poorly** compared to other algorithms in credit scoring task *(Lessmann et al. 2015)*

Therefore, developing a universal profit-driven feature selection framework would contribute to research

Profit Scoring Measures

Measure	Reference	Parameters
R-EMPCS	Garrido et al. 2018	LGD distribution; ROI; shock
EMPCS with data costs	Maldonado et al. 2017	expected LGD; ROI; aquisition costs
EMPCS	Verbraken et al. 2014	LGD distribution; ROI
Expected Costs	Bahnsen et al. 2014	Mean LGD; ROI; loan amount
Profit Contribution	Finlay 2010	LGD, gross payments, profit contribution
EMC	Abdou et al. 2009	ratio of missclassification costs
…	…	…

- Literature suggests different measures to evaluate profitability
- Family of **EMPCS** measures are **state-of-the-art**
- Hence, we will rely on **EMPCS** for feature selection

The EMPCS Measure

Expected **M**aximum **P**rofit for **C**redit **S**coring

- Inspired by *(Verbraken, Verbeke & Baesens, 2013)*
- Applied to credit scoring *(Verbraken et al., 2014)*

Interpretation

incremental profit from applying the scorecard

The EMPCS Measure

		Predicted Class	
		BAD	GOOD
Real Class	BAD		0
	GOOD		0

basic scenario: **all loans are granted**

The EMPCS Measure

		Predicted Class	
		BAD	GOOD
Real Class	BAD	B	0
	GOOD		0

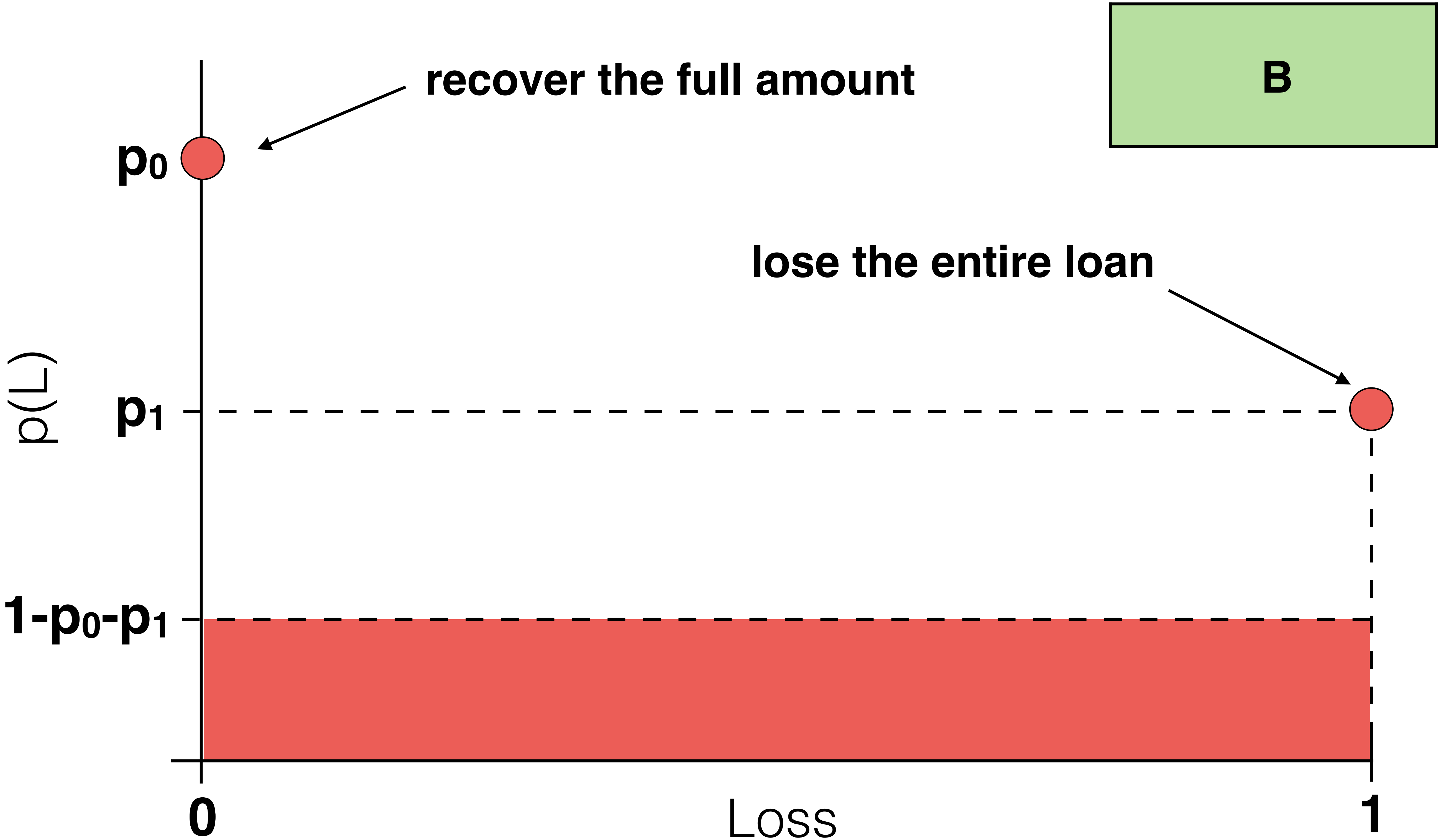
- **B** = expected loss in case of default

The EMPCS Measure

		Predicted Class	
		BAD	**GOOD**
Real Class	**BAD**	**B**	**0**
	GOOD	**-C**	**0**

- **B** = expected loss in case of default
- **C** = return on investment

Loss Distribution



Return on Investment

- $\textbf{ROI} = \dfrac{\textbf{total interest}}{\textbf{principal}}$

- Assumed to be constant for all loans

-C

The EMPCS Measure

$$EMPCS = \int_0^1 \Big[B \cdot \pi_0 F_0(t) - C \cdot \pi_1 F_1(t)\Big] \cdot h(B) d(B)$$

- **B** = expected loss in case of default
- **C** = return on investment

- π_i = prior probabilities of **BAD** and **GOOD**
- $h(B)$ = density function
- $F_i(t)$ = model-based cumulative fractions of **BAD** and **GOOD**
- t = cutoff value

Empirical Experiments

Experiment #1:

- **goal:** check correlation between EMPCS and standard performance measures
- **research question:** is maximizing EMPCS different from optimizing traditional measures?

Experiment #2:

- **goal:** benchmark performance of the EMPCS-maximizing feature selection compared to conventional strategies
- **research question:** does profit-driven feature selection lead to scorecards with higher expected profit?

Data Library

Data:

ten credit scoring data sets

Data Label	Sample Size	Num. Features	Default Rate
australian	690	42	0.4449
german	1,000	61	0.3000
thomas	1,225	28	0.2637
bene1	3,123	83	0.3333
hmeq	5,960	20	0.1995
bene2	7,190	28	0.3000
uk	30,000	51	0.0400
lendingclub	43,344	206	0.1351
pakdd	50,000	373	0.2608
gmsc	150,000	68	0.0668

Experiment #1

- Cross-validate models with **different feature subsets**
- Compute **rank correlations** between **evaluation measures**
- High correlation = measures choose the same feature sets

Experiment #1

- Cross-validate models with **different feature subsets**
- Compute **rank correlations** between **evaluation measures**
- High correlation = measures choose the same feature sets

Measure	AUC ROC	H-measure	Accuracy	Brier Score	EMPCS
AUC ROC	1				
H-measure	0.7111	1			
Accuracy	0.1236	0.3795	1		
Brier Score	-0.6753	-0.8359	-0.3744	1	
EMPCS	0.2349	0.6941	0.4521	-0.6022	1

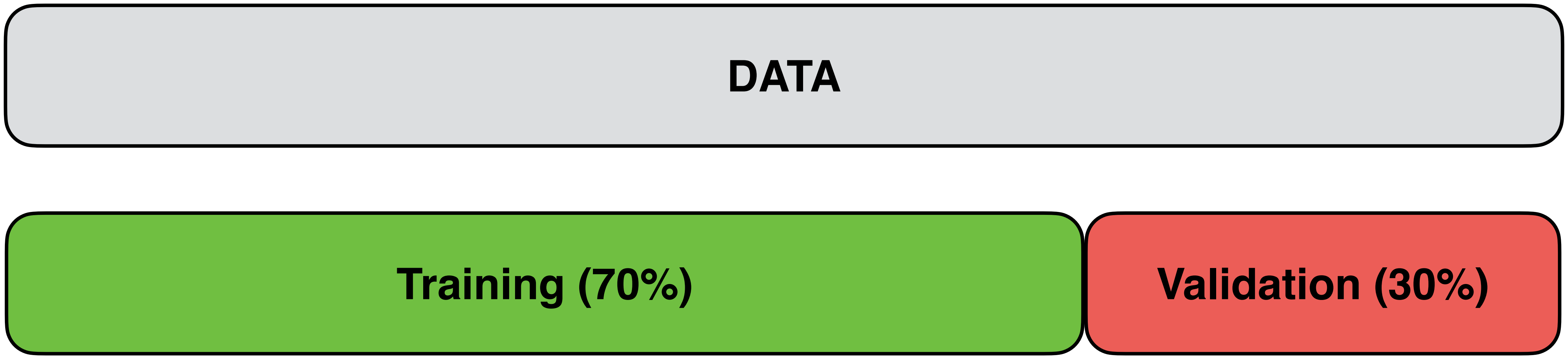
Experiment #1

- Cross-validate models with **different feature subsets**
- Compute **rank correlations** between **evaluation measures**
- High correlation = measures choose the same feature sets

Measure	AUC ROC	H-measure	Accuracy	Brier Score	EMPCS
AUC ROC	1				
H-measure	0.7111	1			
Accuracy	0.1236	0.3795	1		
Brier Score	-0.6753	-0.8359	-0.3744	1	
EMPCS	0.2349	0.6941	0.4521	-0.6022	1

Optimizing EMPCS results in a different ranking of feature subsets compared to standard measures

Experiment #2



Experiment #2

DATA

Training (70%) | Validation (30%)

Feature selection:

- Sequential forward selection
- Sequential backward selection
- Genetic algorithm

Objective function:

- AUC ROC
- EMPCS

$\left.\right\}$ **4-fold CV**

Experiment #2

DATA

Training (70%) | Validation (30%)

Evaluate the models

Feature selection:

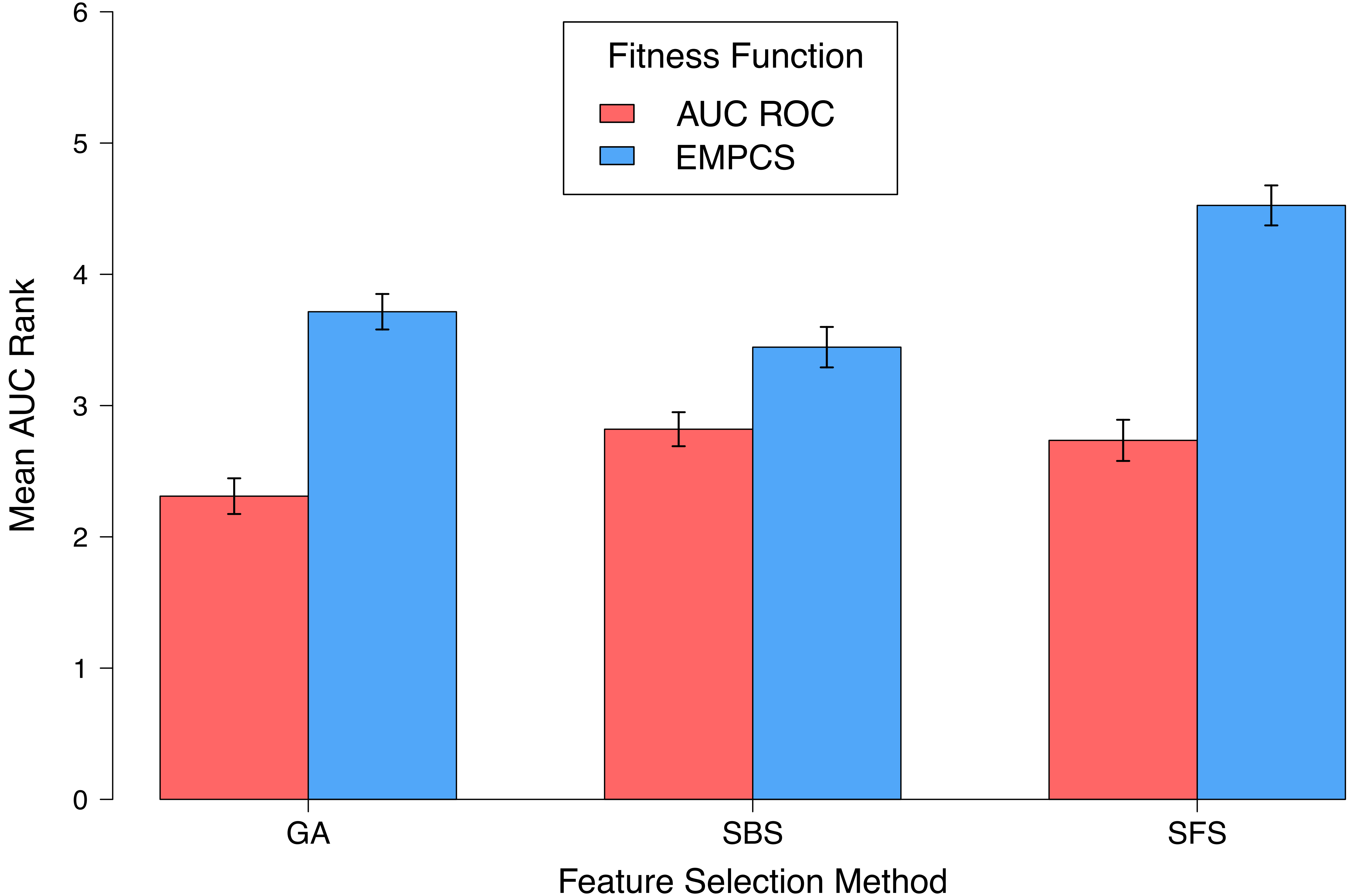
- Sequential forward selection
- Sequential backward selection
- Genetic algorithm

4-fold CV

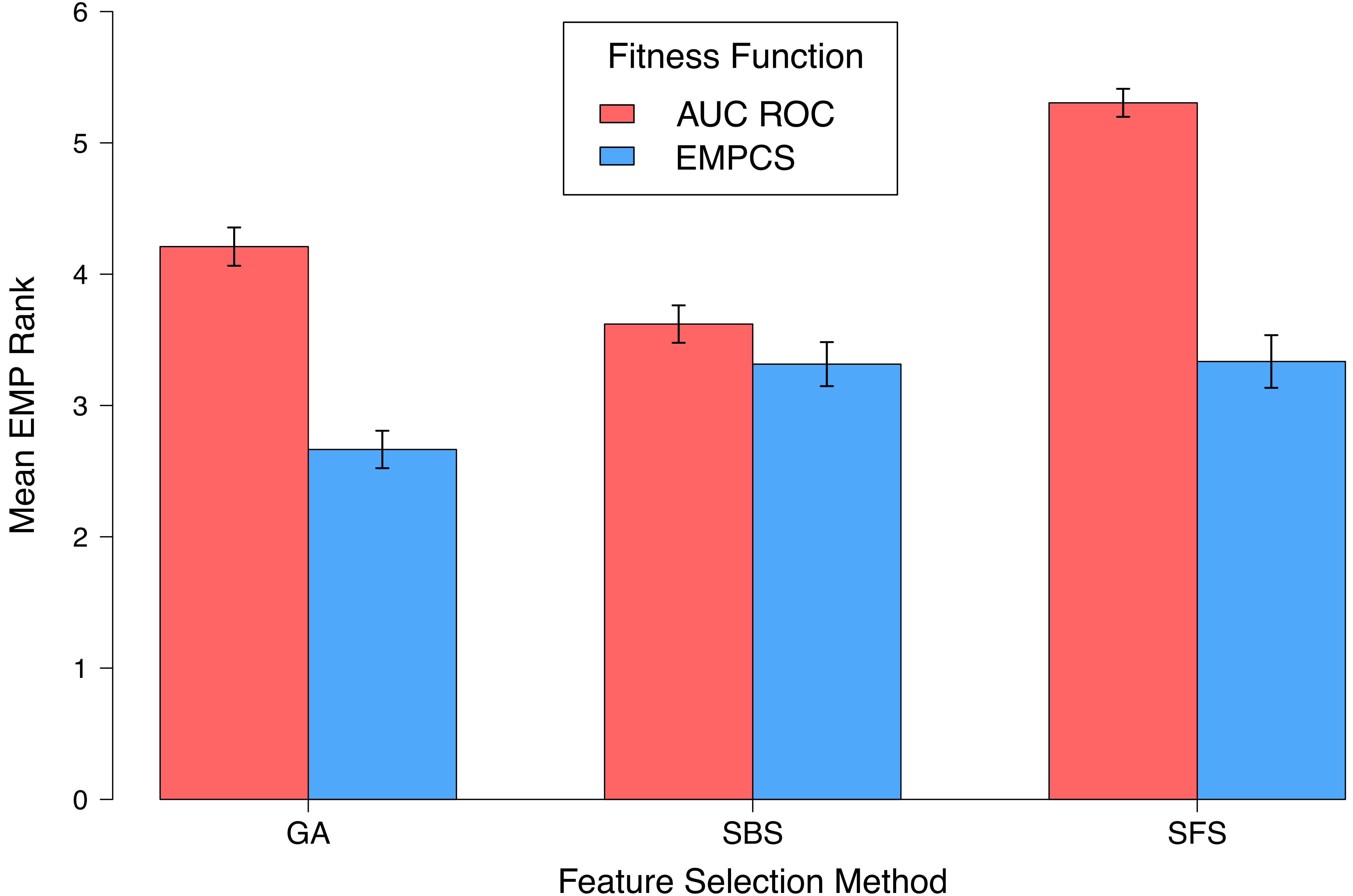
Objective function:

- AUC ROC
- EMPCS

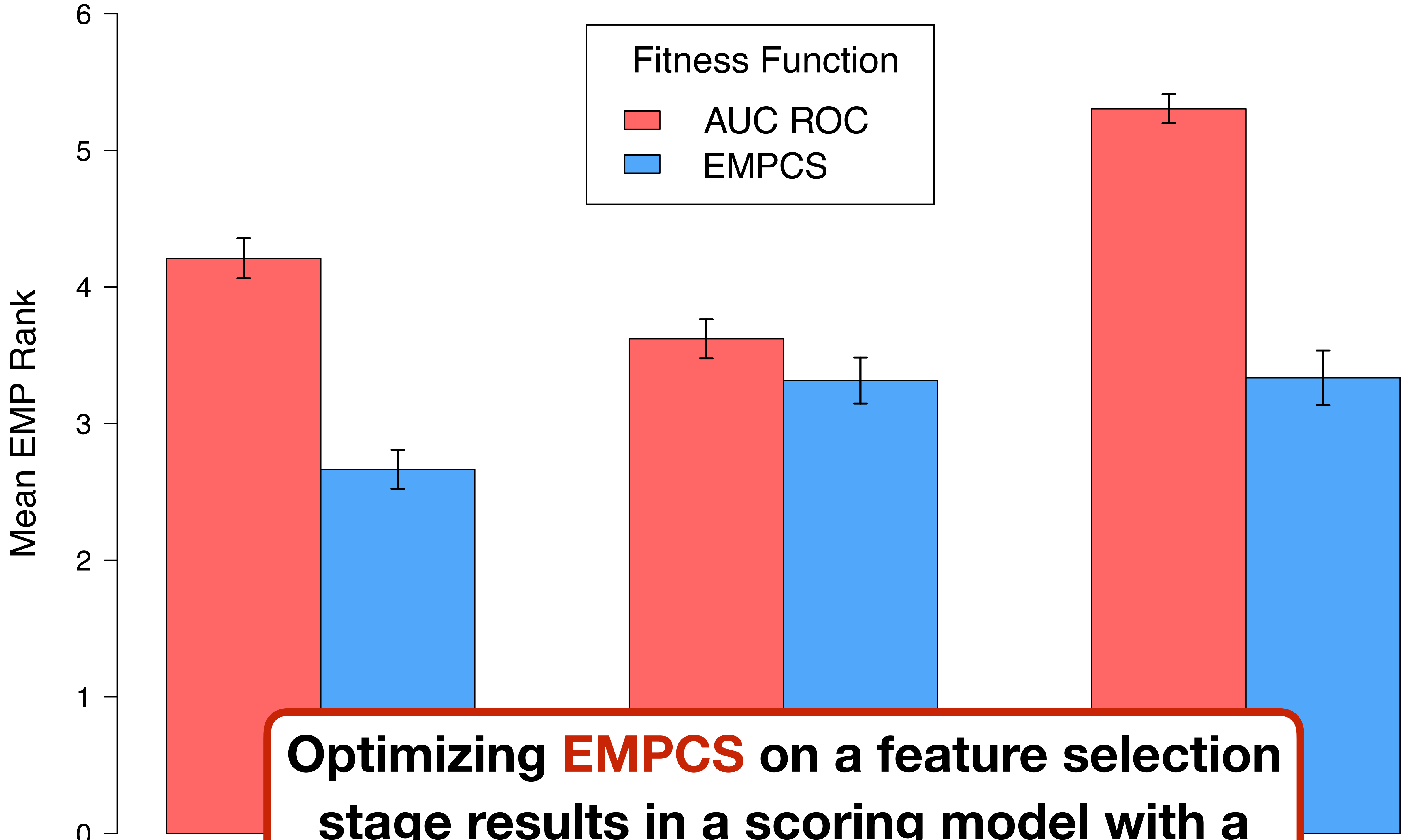
Results: Mean AUC Rank



Results: Mean EMPCS Rank



Results: Mean EMPCS Rank



Optimizing EMPCS on a feature selection stage results in a scoring model with a lower EMPCS rank on holdout data

Conclusions & Next Steps

Conclusions:

- using EMPCS for feature selection **increases the expected profitability** of the scorecards
- results emphasize importance of using **business-inspired indicators** on the **feature selection stage**

Next Steps:

- extending profit-driven framework to **other modeling stages**
- benchmarking **a rich set** of EMPCS-based wrappers and filters
- applying the profit-oriented approach to **other business areas**

Discussion & Questions

Thank you for your attention!