

# Предсказание оттока абонентов с помощью метода опорных векторов

Выполнил: Кравченко Евгений, 344 гр  
Научный руководитель: д.ф.-м.н., проф. А.Н. Терехов

# Цель работы

Рассмотреть применимость классификаторов, основанных на SVM к задаче предсказания оттока абонентов.

Задачи:

- Рассмотреть уже существующие решения данной задачи
- Разработать классификатор, предсказывающий уход абонентов
- Рассмотреть возможности ансамблей с использованием SVM
- Оптимизировать классификатор для данной задачи

# Метрики

$$\textit{precision} = \frac{TP}{TP+FP}$$

$$AUC = \int_0^1 \frac{TP}{P} d\frac{FP}{N}$$

$$\textit{recall} = \frac{TP}{P}$$

$$F_1 = \frac{2 \times \textit{precision} \times \textit{recall}}{\textit{precision} + \textit{recall}}$$

# Обзор

Лучший результат предыдущих лет:  
бустинг над решающими деревьями

AUC	Precision	Recall	F1
0.92	0.75	0.72	0.73

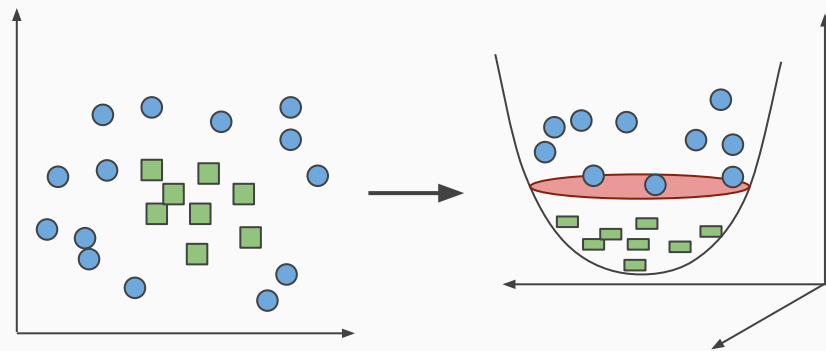
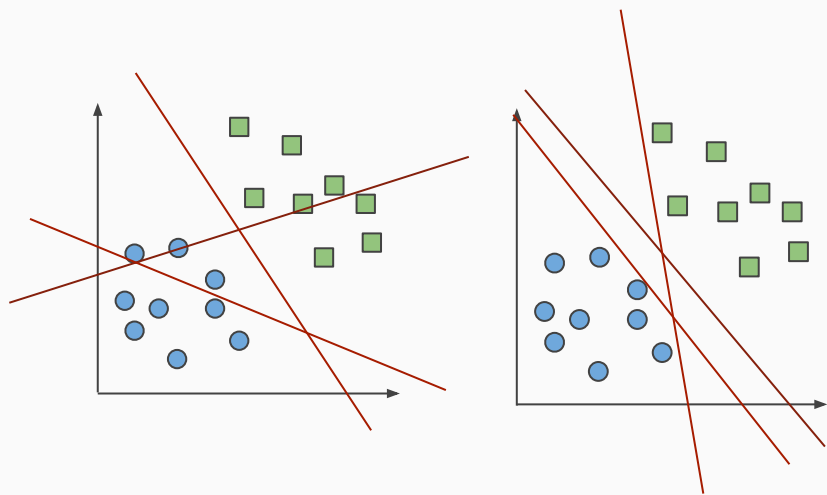
Другие перспективные модели: SVM, bagging с SVM

# Данные

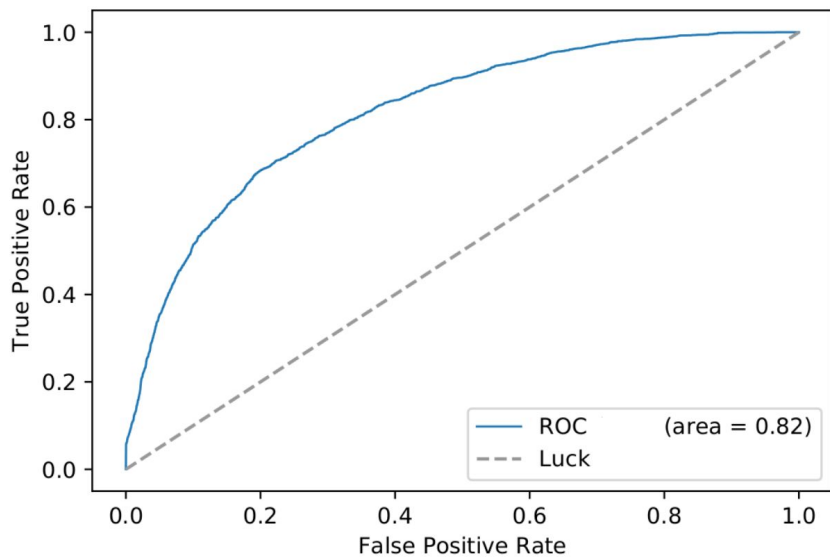
Данные по 300000 клиентам за 4 месяца:

- количество входящих/исходящих звонков и СМС
- количество использованного интернет-трафика
- количество обращений в справочные службы
- личная информация

# SVM

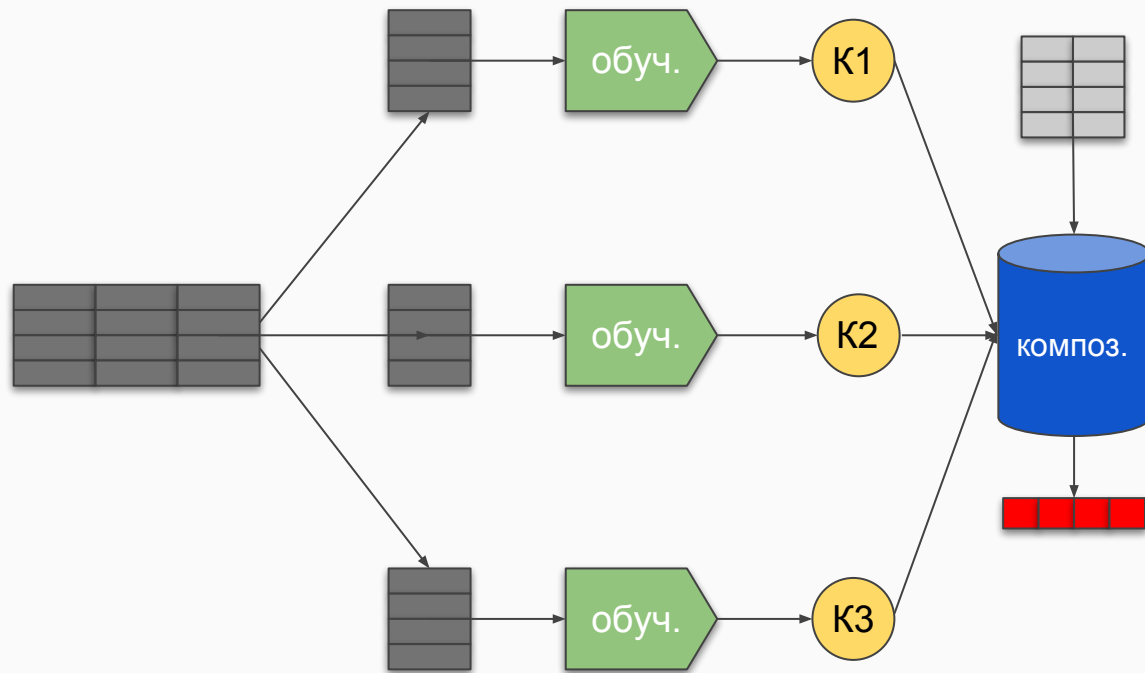


# SVM



AUC	Precision	Recall	F1
0.82	0.58	0.43	0.49

# БЭГГИНГ





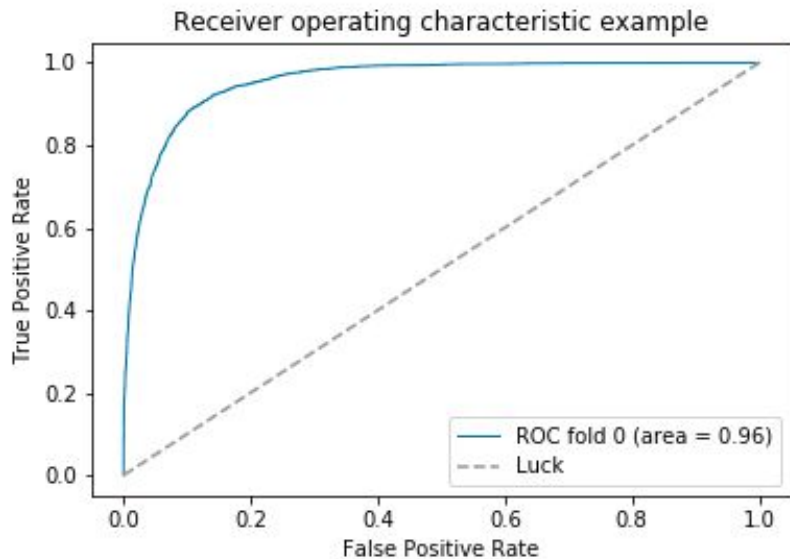
# Подбор параметров

RBF - ядро:

$$K(x, x') = e^{-\gamma \|x - x'\|^2}$$

$\gamma \backslash C$	$10^{-3}$	$10^{-1}$	1	$10^3$	$10^5$
$10^{-3}$	0.92	0.93	0.93	0.93	0.91
$10^{-1}$	0.95	0.96	0.95	0.95	0.94
1	0.89	0.91	0.91	0.93	0.92
$10^3$	0.94	0.93	0.94	0.92	0.94
$10^5$	0.93	0.93	0.95	0.92	0.92

# Лучший результат



AUC	Precision	Recall	F1
0.96	0.76	0.77	0.76

# Результаты

	Precision	Recall	AUC
Лучший результат прошлых лет	0.75	0.72	0.92
SVM	0.58	0.43	0.81
Бэггинг (SVM)	0.76	0.77	0.96

Можно сделать вывод, что классификаторы, основанные на ансамблях из SVM, применимы в задаче предсказания оттока абонентов.