

# Модель ML для прогноза шизофрении

- Бойзода Хаким



# Введение



- - Шизофрения — серьёзное психическое расстройство.
- - Ранняя диагностика улучшает лечение.
- - Машинное обучение помогает выявить болезнь на ранней стадии.

- ▶ - Разработка модели для прогнозирования шизофрении.
- ▶ - Использование исторических медицинских данных.
- ▶ - Оценка эффективности модели.

## ЦЕЛЬ ПРОЕКТА

# Обзор набора данных

- - Источник : [Kaggle.com/datasets](https://www.kaggle.com/datasets)
- - количество записей: 10,000
- - признаки: 13.
- - Основные признаки: gender, income\_lvl, live\_area, Substance\_use, Suicide\_Attempt, Stress\_Factors.
- - Предварительная обработка: пропущенные значения, кодирование.
- - Распределение классов: 0 – 7113, 1 - 2887

# Выбор и обучение модели

- - Используемые алгоритмы:
- - Разделение данных на обучающую и тестовую выборки:  
train - 85%, test – 15%.
- - Метрики: Roc-Auc, precision, recall, F1-score:

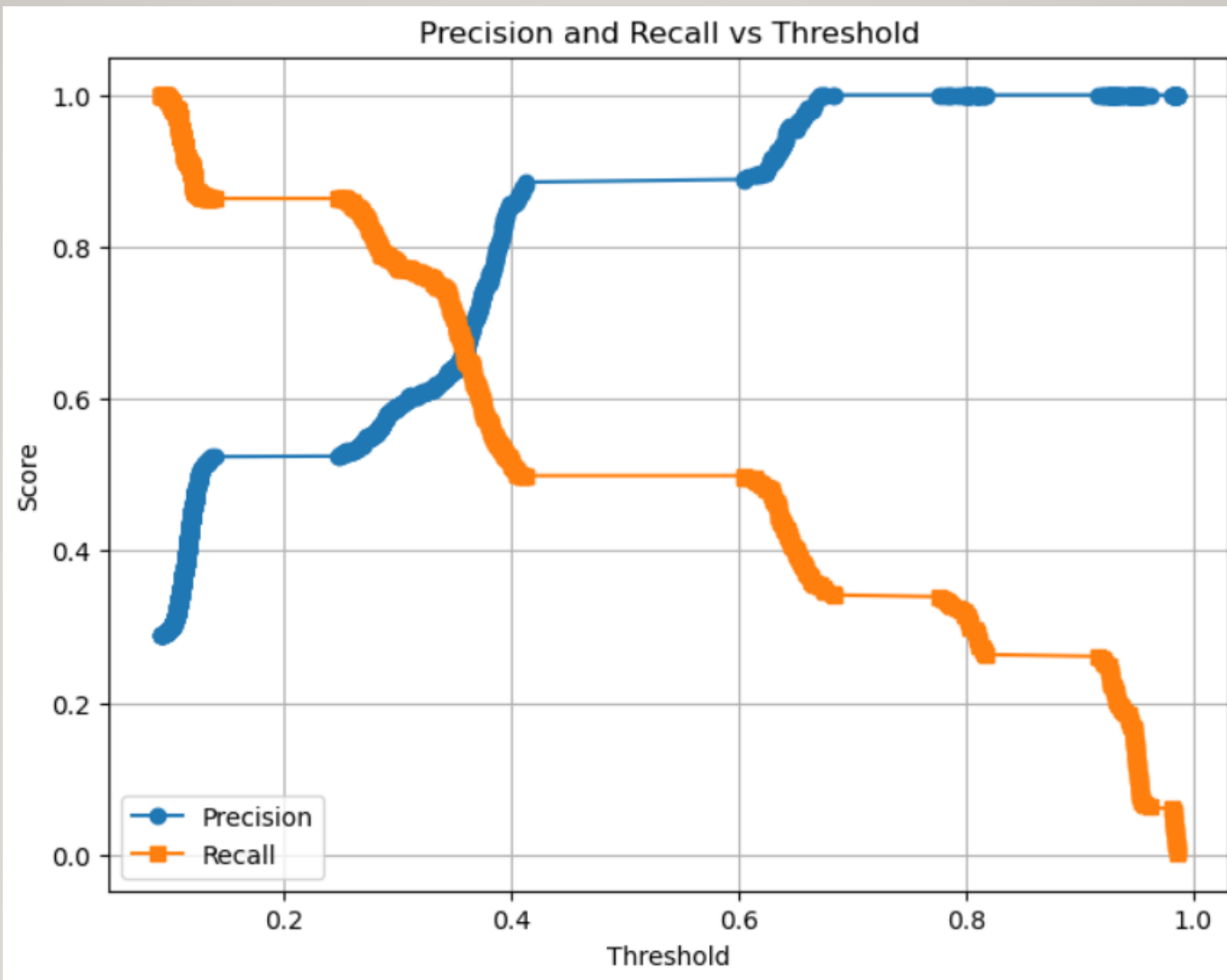
	Model	ROC-AUC	Recall	Precision	F1 Score
0	LogisticRegression	0.825161	0.473914	0.871727	0.613724
1	DecisionTreeClassifier	0.677137	0.559485	0.525820	0.541939
2	RandomForestClassifier	0.799923	0.486950	0.784686	0.600566
3	KNeighborsClassifier	0.668953	0.246539	0.631759	0.354481
4	CatBoostClassifier	0.812570	0.497548	0.793090	0.611032

# Результаты и анализ

- Самой лучшей моделью оказалась логистическая регрессия с метриками:  
Roc-Auc: 0.82, recall: 0.47, precision: 0.87, F1: 0.61
- Новая цель: Подбор наилучших гиперпараметров и улучшение модели

# Улучшения

- Новые гиперпараметры:  
C: 0.03, penalty: 'l2', solver: 'saga', max\_iter: 120
- Новая цель найти оптимальный порог.



Самый оптимальный порог: 0.35



# Окончательный результат

- Модель: Логистическая регрессия
- Гиперпараметры: C: 0.03, penalty: 'l2', solver: 'saga', max\_iter: 120
- Порог: 0.35

Metrics	Scores
Roc-Auc	0.825
Precision	0.641
Recall	0.706
F1-score	0.672

# Использование

- Перейдём к демонстрации модели в Streamlit 