

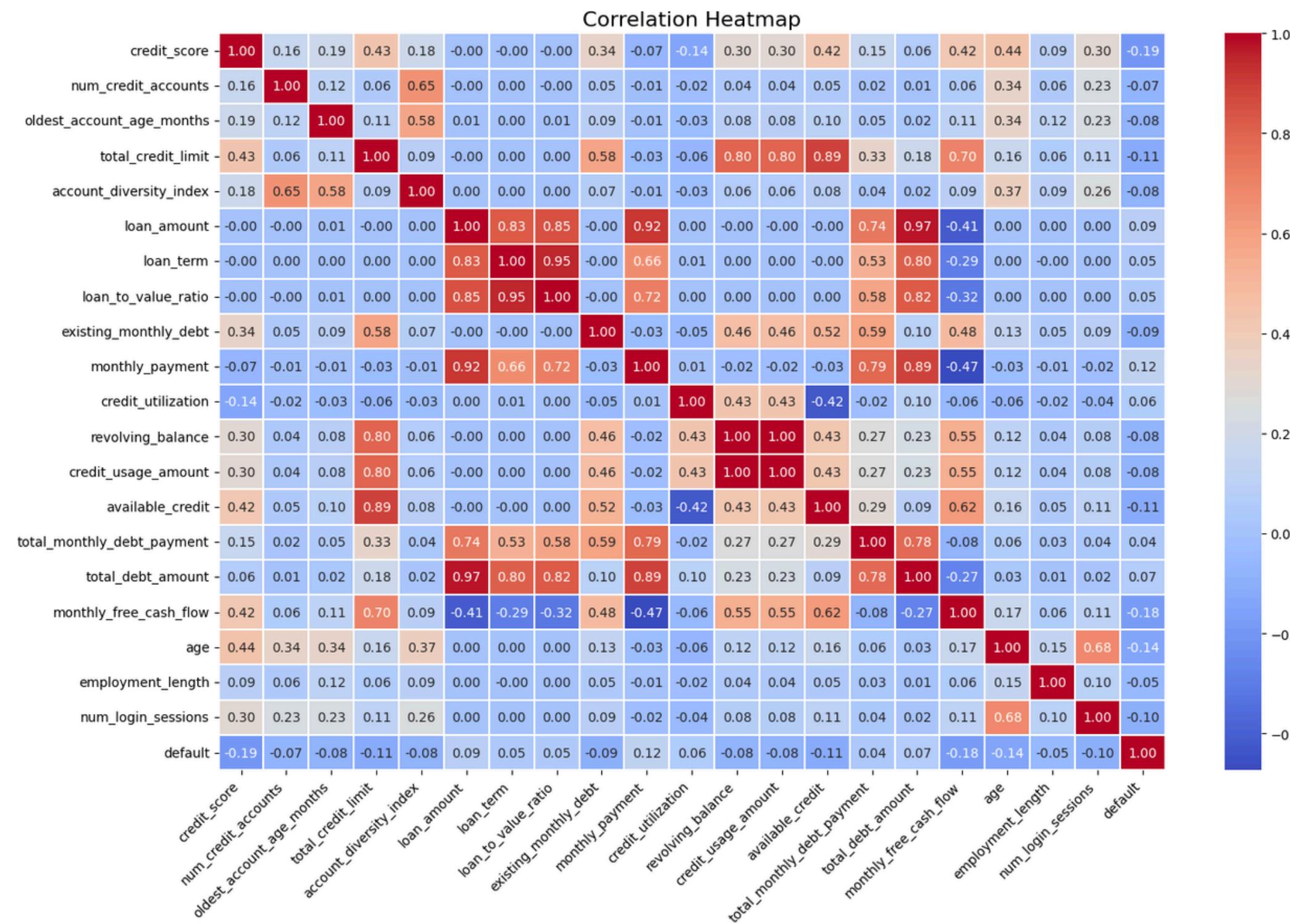
Прогнозирование кредитных дефолтов

Команда: reze

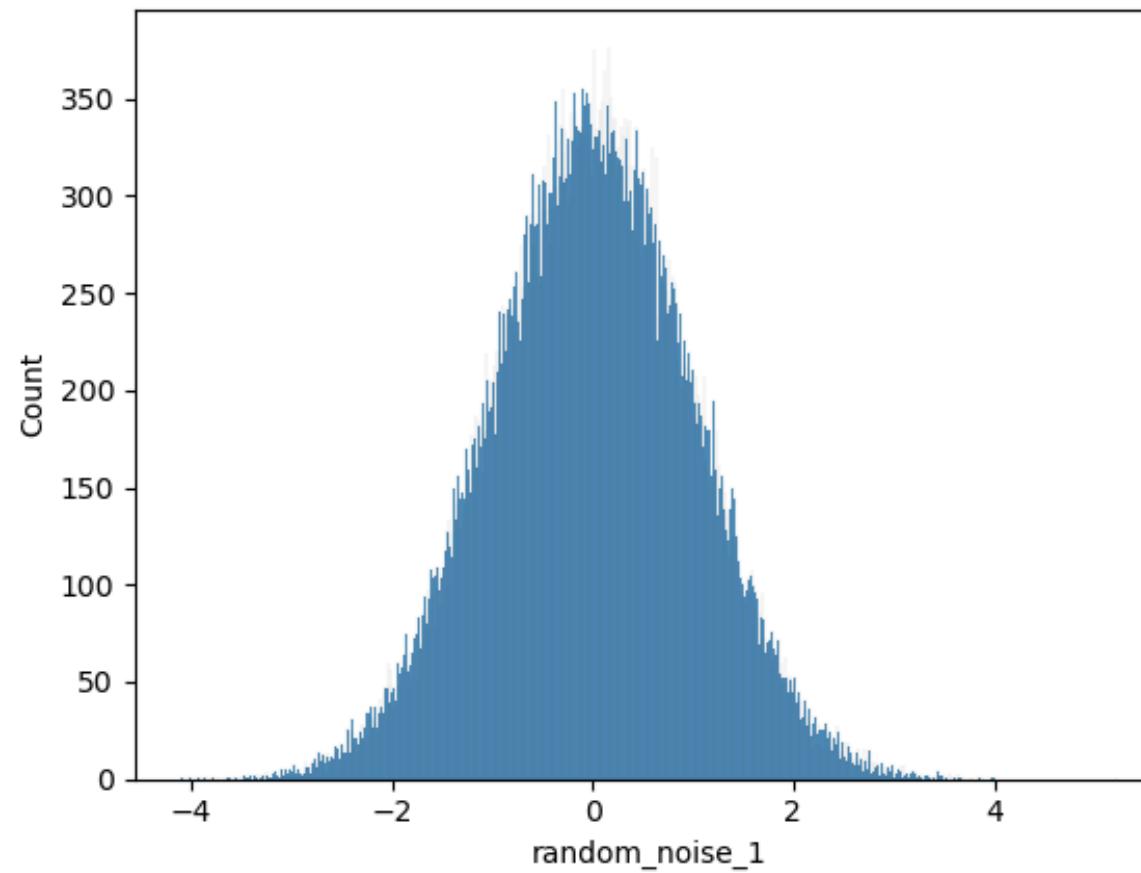
Трудности

- Координация
- Разделение задач
- Нехватка знаний в области

Матрица корреляции



ЧТО МЫ НАКОПАЛИ



```
1 merged['employment_type'].unique()
✓ 0.0s

array(['Full-time', 'FULL_TIME', 'Full Time', 'Fulltime', 'Part Time',
       'Self Employed', 'FT', 'Self Emp', 'Contractor', 'Self-employed',
       'SELF_EMPLOYED', 'Contract', 'PART_TIME', 'CONTRACT', 'PT',
       'Part-time'], dtype=object)
```

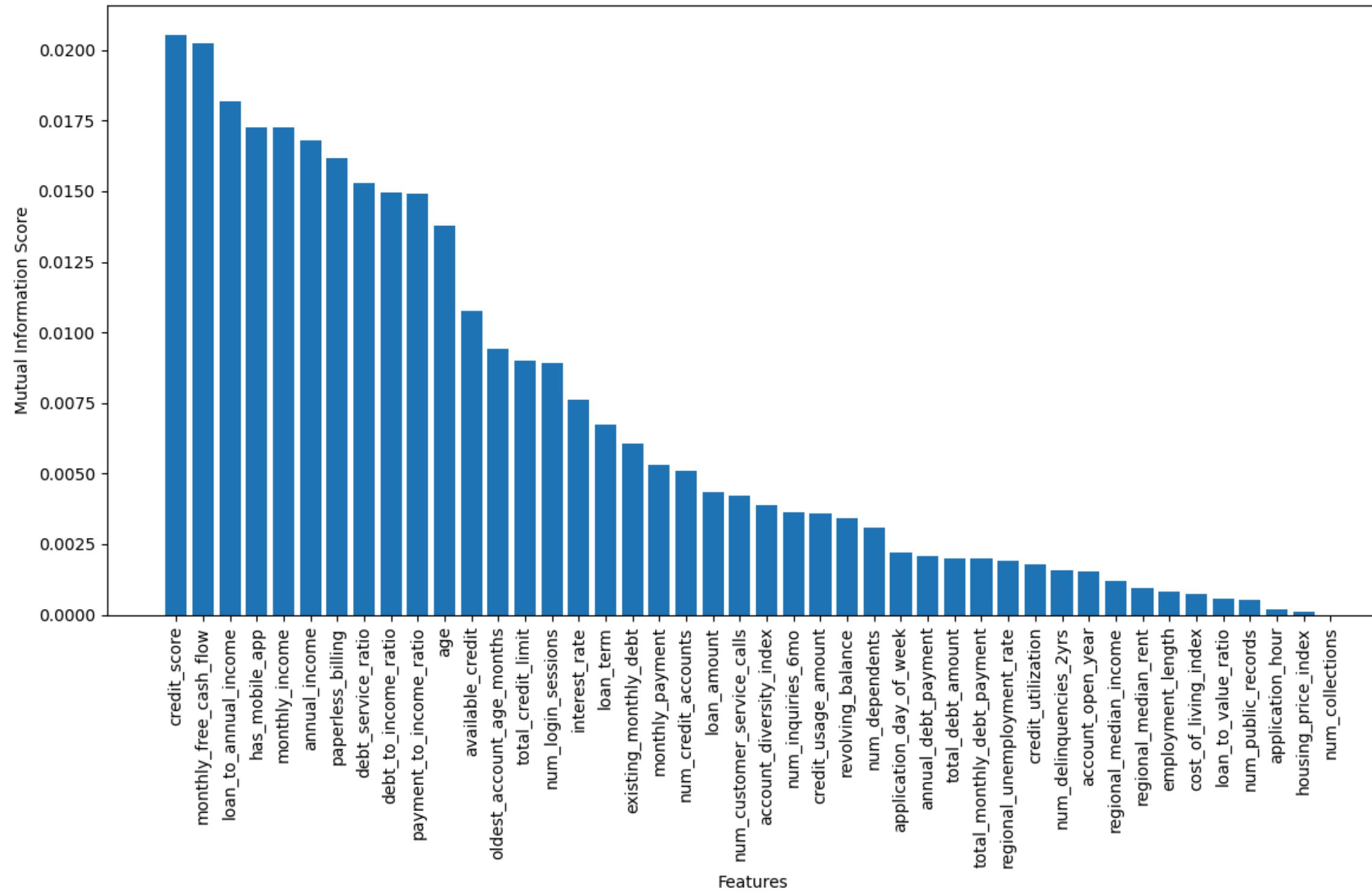
```
1 merged['default'].value_counts(normalize=True)
✓ 0.0s

default
0    0.948955
1    0.051045
Name: proportion, dtype: float64
```

Проделанная работа

- Обработали 2000+ пропущенных записей
- Добавили новые фичи путем скрещивания существующих фичей
- Проверили Logreg, XGBoost, LightGBM, Catboost - победил Catboost

ЧТО МЫ ПОНЯЛИ



Результат

В конечном получили результат: AUC = 0.83