

Прогнозирование результатов футбольных матчей

Создание пользовательского дашборда



Executive Summary

Задача: Создание системы прогноза результатов матчей футбольных команд



Набор данных:
11 европейских высших футбольных дивизионов
Матчи сезонов 2008/2009 - 2015/2016

Набор инструментов:
модель машинного обучения Random Forest
BI-система Tableau



Итоговый продукт:
Разработана интерактивная система прогноза результатов матчей футбольных команд на основе модели "Random Forest" с точностью 70%

Развитие продукта:

Повышение точности прогноза за счет включения наборов данных

- о трансферных переходах игроков
- о характеристиках игроков

Описание проекта

Проект позволит предсказывать результаты футбольных матчей, оптимизируя процесс принятия решений для болельщиков и любителей футбола.



Ключевые шаги для создания продукта

Отбор исходных данных для анализа

Выбор оптимальной модели для прогноза (Random Forest)

Проведение первых итераций расчетов

Создание и проверка первоначальной модели

Feature importance analysis характеристики команды

Улучшение модели за счет кумулятивной статистики команд

Построение новых характеристик команд (рейтинг атаки, защиты, команды и форма команды)

Проведение итоговых расчетов: повышение точности модели

Расчет дополнительной статистики команд (голы и место в лиге)

Преимущества конечного продукта



Интерактивность



Возможность фильтрации
по странам и командам



Доступность информации за счет
графического представления
данных



Характеристики команд,
основанных на
исторических данных

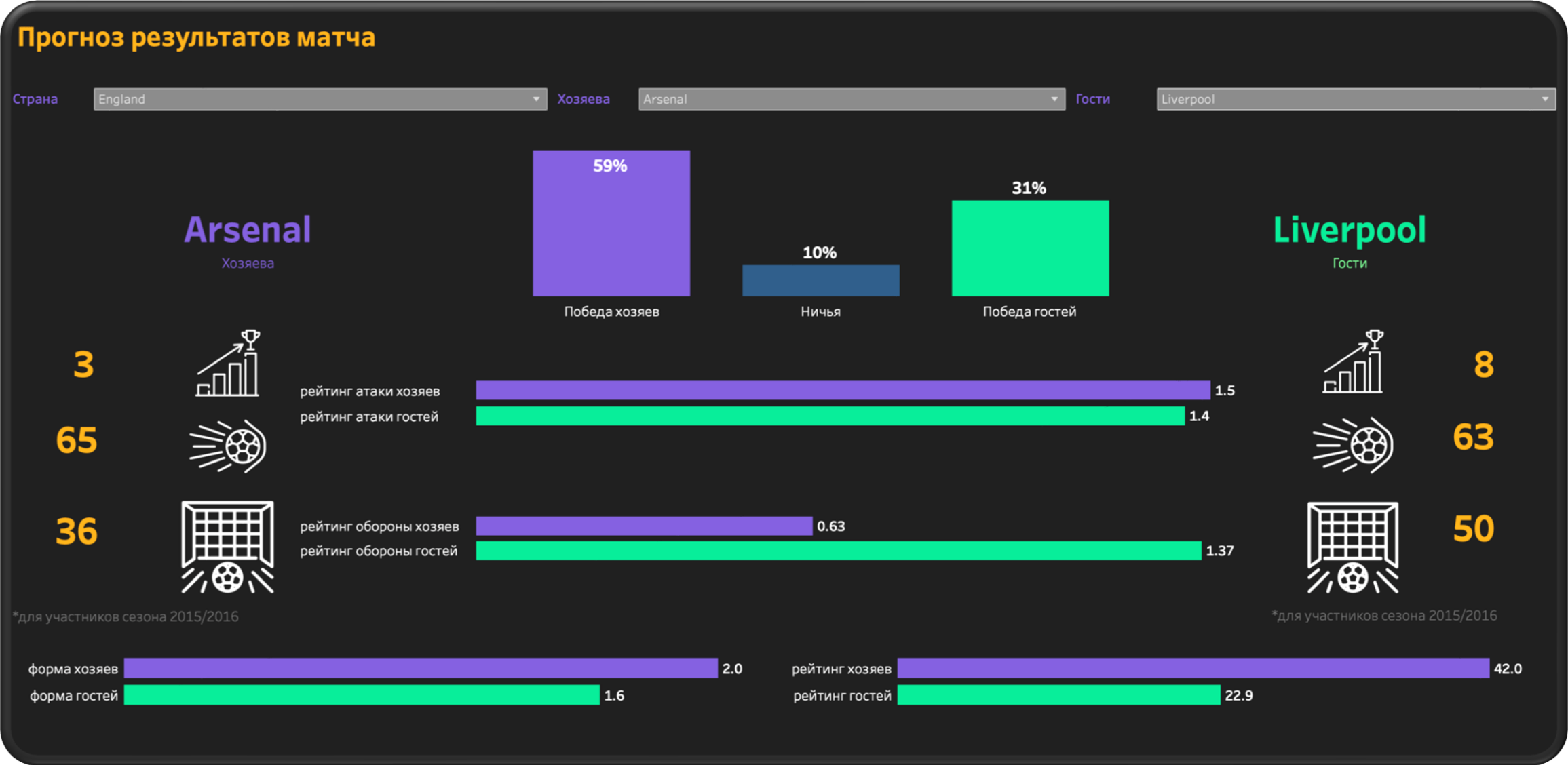


Наличия вероятностей
предсказания



Поддержка принятия решений

Визуализация результатов



Наша команда

Богданова
Дарья

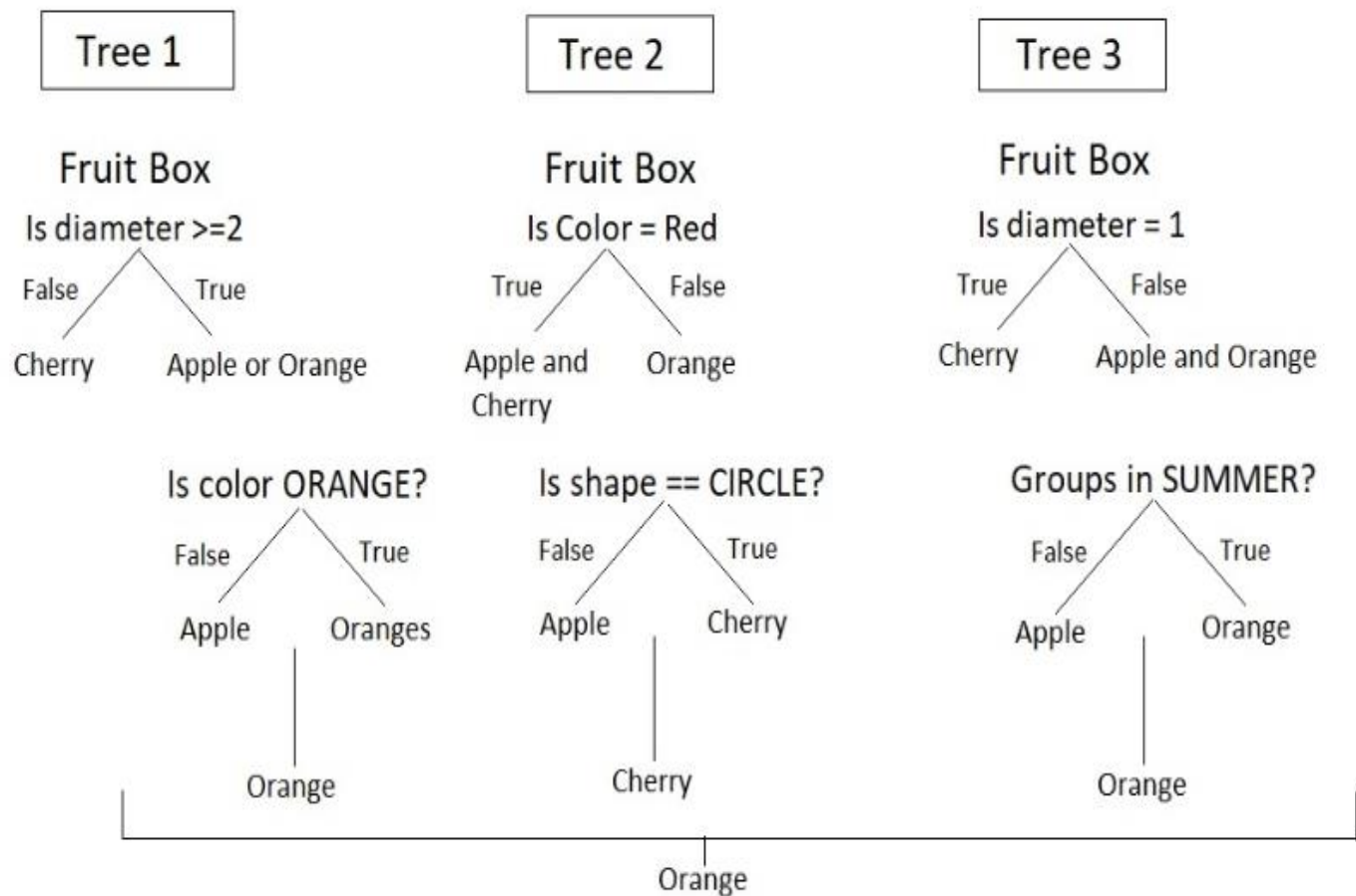
Иванова
Наталья

Забитов
Уллубий

Зарубин
Андрей

Рыкунова
Евгения

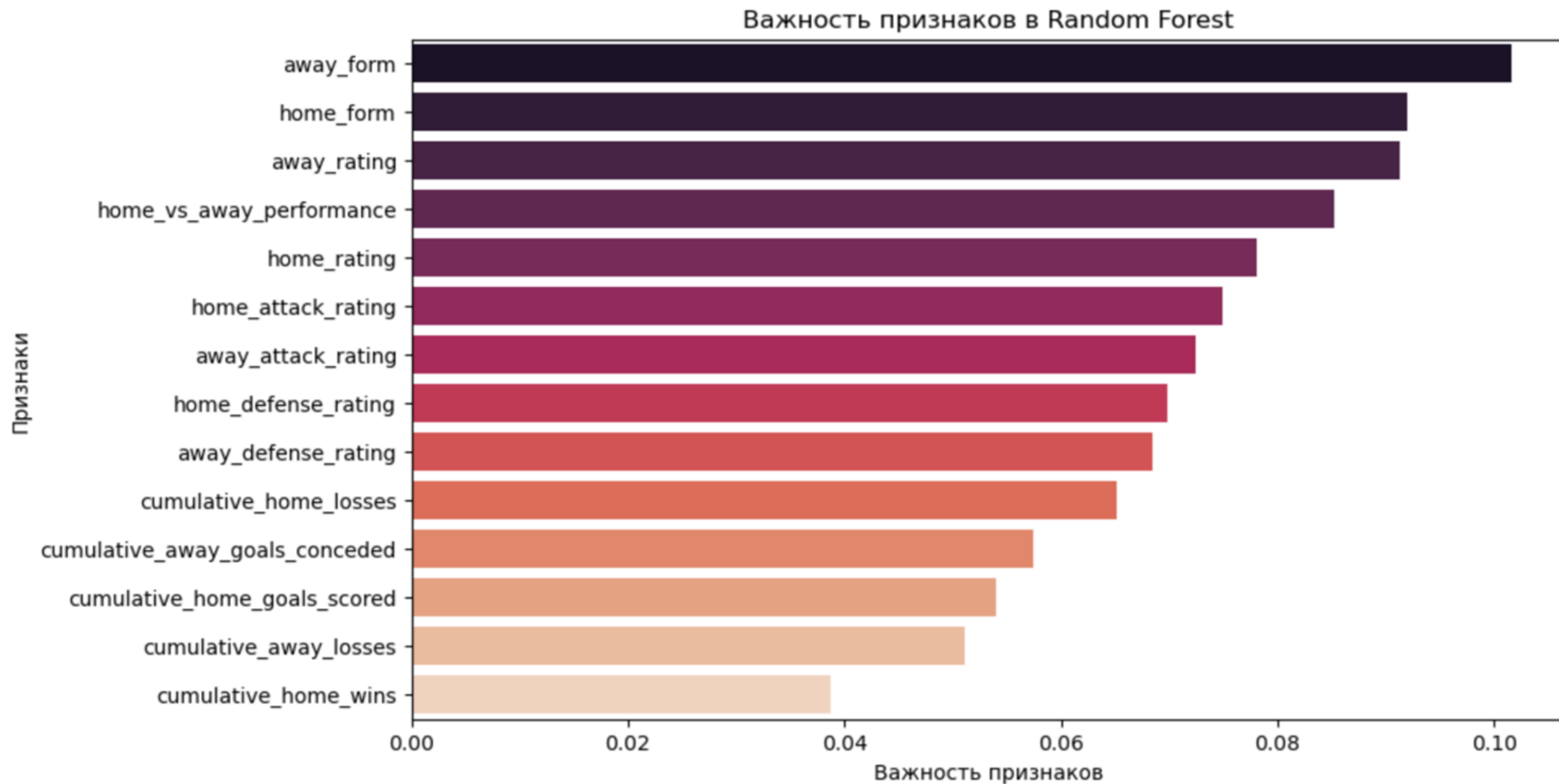
Приложение 1: пример визуализации модели Random Forest



Преимущества данной модели:

- Обработка множества факторов
- Устойчивости к шумам
- Устойчивость к переобучению
- Адаптации к сложным данным
- Устойчивости к небольшим выборкам
- Прогнозирование редких событий

Приложение 2: Feature importance анализ



Наиболее важными критериями для создания дерева решений являются рассчитанные рейтинги, которые будут использоваться для формирования дашборда.

Приложение 3: Используемые формулы для расчета рейтингов команд

$$\text{Рейтинг атаки} = \frac{\text{кол-во забитых голов}}{\text{кол-во матчей}}$$

$$\text{Рейтинг защиты} = \frac{\text{кол-во пропущенных голов}}{\text{кол-во матчей}}$$

$$\text{Рейтинг команды} = \text{победа} \times 3 + \text{ничья} \times 1 + (\text{забитые голы} - \text{пропущенные голы})$$

$$\text{Форма команды} = \frac{\begin{array}{l} \text{сумма набранных очков} \\ \text{за последние 10 игр} \end{array}}{10 \text{ игр}}$$