

Проверка самостоятельности студентов онлайн-курсов

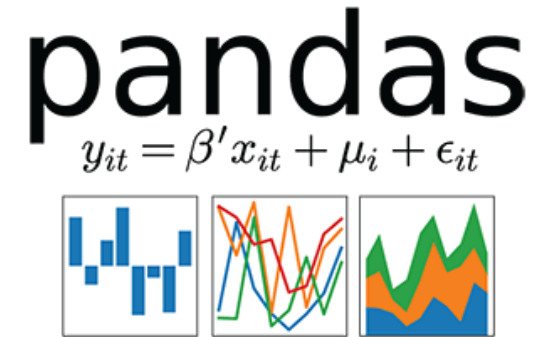
Камкова Екатерина
244 группа
24.05.2019

Цель: используя логи онлайн-курса выявить студентов, подозрительных на списывание

- Собрать из логов необходимые данные и организовать в формат, подходящий для машинного обучения
- Подобрать подходящую задачу модель машинного обучения
- Реализовать простой консольный интерфейс для взаимодействия пользователя и программы



Используемые технологии



Аналогичные исследования:

- Dropout Prediction in MOOCs: Using Deep Learning for Personalized Intervention
- MOOC Performance Prediction via Clickstream Data and Social Learning Networks
- Your Click Decides Your Fate: Inferring Information Processing and Attrition Behavior from MOOC Video Clickstream Interactions
- Big data analytics for MOOC video watching behavior based on Spark

Работа с логами

Необходимые данные:

- Количество начатых задач
- Количество решённых задач
- Среднее количество попыток, необходимое студенту
- Количество просмотренных видео-уроков
- Сколько раз в среднем студент пересматривает видео
- Количество запросов на правильный ответ
- Участие в обсуждениях на форуме
- И другие ...

Модель

- Бинарная классификация
- Обучение с учителем

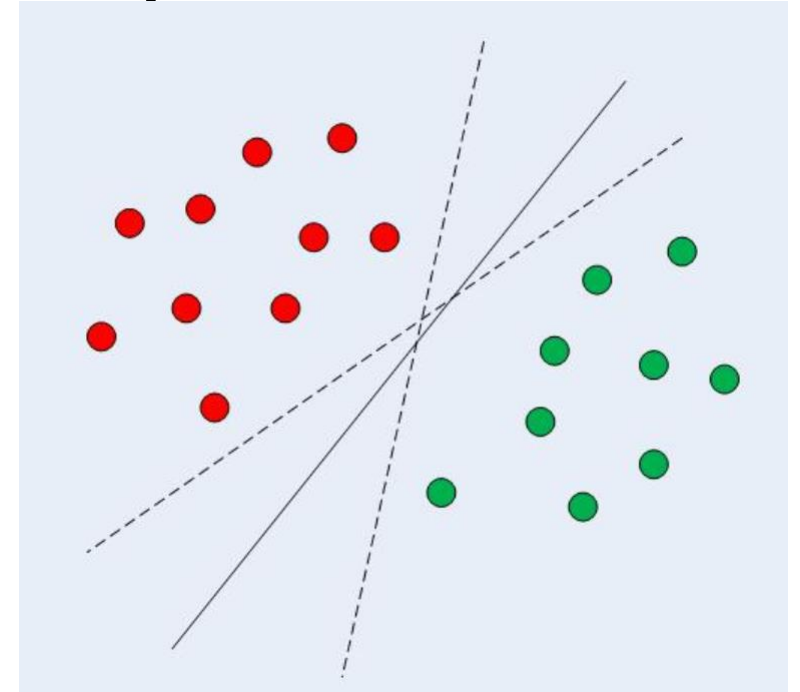
Рассматриваемые модели:

1. Метод опорных векторов
2. Дерево принятия решений

Метод опорных векторов

Основная идея:

- Гиперплоскость
- Расстояние между гиперплоскостью и ближайшими точками данных



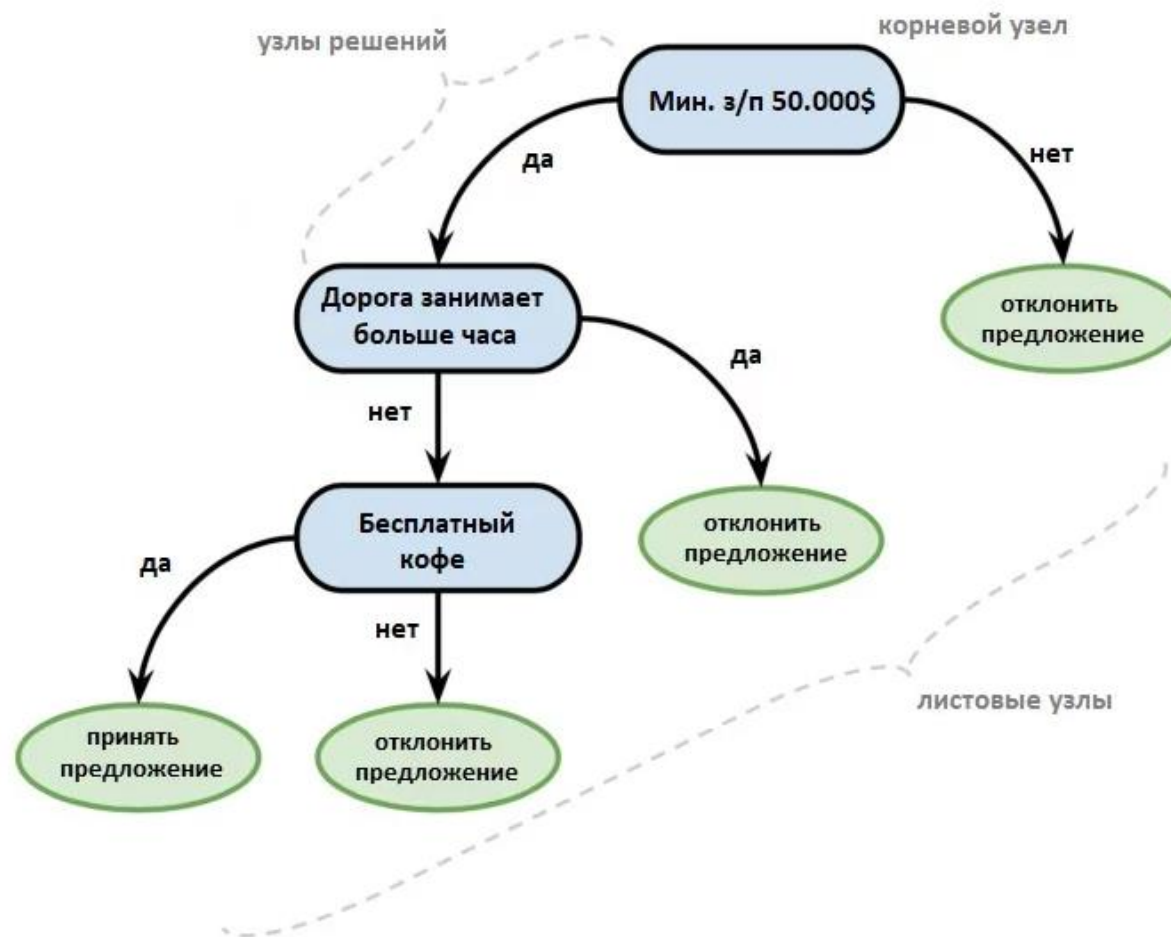
Точность: 0.665

Дерево принятий решений

Основная идея:

- Узел – входная переменная
- Лист – выходная переменная

Точность предсказания – 0.91



Интерфейс

1. Создание обучающей таблицы
 - Загрузить уже существующую
 - Создать из логов и таблицы ответов
2. Создание тестируемой таблицы
 - Загрузить уже существующую таблицу без результатов
 - Создать из логов
3. Обучение модели и предсказание самостоятельности студентов, сохранение результатов
4. Добавление проверенных студентов в обучающую таблицу

Отдельно прописаны проверки моделей SVM и дерева принятий решений

Итоги

- Реализован алгоритм, извлекающий необходимую информацию
- Выбрана модель машинного обучения
- Реализована классификация учеников
- Реализован базовый интерфейс

Ссылка на репозиторий:

<https://github.com/katerina-kamkova/CheatCheck>