\documentclass[12pt]{article}

\usepackage[utf8]{inputenc}

\usepackage[russian]{babel} % Включаем пакет для поддержки русского языка

\usepackage[left=2cm,right=2cm,

top=2cm,bottom=2cm,bindingoffset=0cm]{geometry}

\begin{document}

\thispagestyle{empty}

\begin{center}

\textbf{МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ \\\*РОССИЙСКОЙ ФЕДЕРАЦИИ}

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ

УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ\\\* \textbf{«БЕЛГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ\\\* ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ им. В.Г.ШУХОВА»\\\* (БГТУ им. В.Г.Шухова)}\\\*\vspace{3cm}

Кафедра программного обеспечения вычислительной техники и\\\* автоматизированных систем

\\\*\vspace{3cm} Отчёт по компьютерной практике

\end{center}

\\\*\vspace{1cm}

\begin{flushright}

Выполнил: ст. группы ВТ-12 \\\*\vspace{0.2cm}Воскобойников Илья Срегеевич\\\*\vspace{0.5cm}

\line (3,0) {5 cm}\\\*

(подпись студента)\\\*

\vspace{0.2cm}

Проверил: Кабалянц Пётр Степанович\\\*\vspace{0.5cm}

\line (3,0) {5 cm}\\\*

(подпись руководителя практики)\\\*\vspace{0.5cm}

Оценка\line (3,0) {5 cm}

\end{flushright}

\begin{center}

\vspace{\stretch{1}}

Белгород 2019

\end{center}

\newpage

\begin{center}

\Large\textbf{Задание к работе}

\end{center}

\begin{large}

\vspace{0.5cm}

\textbf{Тема: Реализация алгоритма для прохождения игры Gold Train}\\\*

\\\*

Уточнение задания:\\\*

\begin{itemize}

\item Чтение с картинки

\item Есть река

\item Уровень >40

\item Любое стройство

\item Вывод графического решения

\end{itemize}

\end{large}

\newpage

\begin{center}

\Large\textbf{Изучение и анализ предметной области}

\end{center}

\vspace{1cm}

\\\*

\begin{large}

Задача расспознания с картики будет решаться с помощю бибилиотеки OpenCV.

\\\*

\\\*

OpenCV (англ. Open Source Computer Vision Library, библиотека компьютерного зрения с открытым исходным кодом) — библиотека алгоритмов компьютерного зрения, обработки изображений и численных алгоритмов общего назначения с открытым кодом. Реализована на C/C++.Может свободно использоваться в академических и коммерческих целях — распространяется в условиях лицензии BSD.

\end{large}

\newpage

\begin{center}

\Large\textbf{Разработка общего алгоритма решения задачи}

\vspace{0.5cm}

\end{center}

\begin{large}

В самом начале нам надо считать всю нужную информацию из файла пометить ее в DataFrame. После мы применяем корелляцию.\\\*

Корреляция — статистическая взаимосвязь двух или более случайных величин (либо величин, которые можно с некоторой допустимой степенью точности считать таковыми). При этом изменения значений одной или нескольких из этих величин сопутствуют систематическому изменению значений другой или других величин. Математической мерой корреляции двух случайных величин служит корреляционное отношение n либо коэффициент корреляции R.\\\*

Далее мы применяем следующие методы решения:

\section{KNN Классификация}

KNN расшифровывается как k Nearest Neighbor или k Ближайших Соседей — это один из самых простых алгоритмов классификации, также иногда используемый в задачах регрессии. \\\*

Задача классификации в машинном обучении — это задача отнесения объекта к одному из заранее определенных классов на основании его формализованных признаков. Каждый из объектов в этой задаче представляется в виде вектора в N-мерном пространстве, каждое измерение в котором представляет собой описание одного из признаков объекта.\\\*

Для обучения классификатора необходимо иметь набор объектов, для которых заранее определены классы. Это множество называется обучающей выборкой, её разметка производится вручную.

\begin{center}

\textbf{Алгоритм}\\\*

\end{center}

Для классификации каждого из объектов тестовой выборки необходимо последовательно выполнить следующие операции:\\\*

\begin{itemize}

\item Вычислить расстояние до каждого из объектов обучающей выборки

\item Отобрать k объектов обучающей выборки, расстояние до которых минимально

\item Класс классифицируемого объекта — это класс, наиболее часто встречающийся среди k ближайших соседей

\end{itemize}

\section{Логистическая регрессия}

Логистическая регрессия или логит-регрессия — это статистическая модель, используемая для прогнозирования вероятности возникновения некоторого события путём подгонки данных к логистической кривой.\\\*

Логистическая регрессия обычно используется для оценки вероятности того, что образец принадлежит к определенному классу. Если оценочная вероятность больше 50\%, тогда модель прогнозирует, что образец принадлежит к данному классу ( называемому положительным классом, помеченным “1”), а иначе — что не принадлежит (т.е. относится к отрицательному классу, помеченному “О”). Это делает ее двоичным классификатором.

\section{Lasso регрессия}

Метод регрессии “лассо” (LASSO, Least Absolute Shrinkage and Selection Operator) заключается во введении дополнительного слагаемого регуляризации в функционал оптимизации модели, что часто позволяет получать более устойчивое решение. Условие минимизации квадратов ошибки при оценке параметров B выражается следующей формулой:\\\*

\\\*

\hat{\beta} = \arg \min \left(\sum\_{i=1}^n (y\_i - \sum\_{j=1}^m \beta\_j x\_{ij})^2 + \lambda |\boldsymbol{\beta}|\right)

\section{KNN регрессия}

Ключевая идея, лежащая в основе метода k ближайших соседей, состоит в формулировке модели в терминах евклидовых расстояний в исходном многомерном пространстве признаков. Задача заключается в том, чтобы для каждой тестируемой точки $x\_0$ найти такую $\delta$ - окрестность (многомерный эллипсоид), чтобы в ней поместилось k точек с известными значениями y . Тогда прогноз $f(\boldsymbol{x}\_0)$ можно получить, усредняя значения отклика всех обучающих наблюдений из $\delta$.\\\*

В случае использования метода для регрессии, объекту присваивается среднее значение по k ближайшим к нему объектам, значения которых уже известны.

\end{large}

\end{document}