**1 СЛАЙД:**

Здравствуйте, меня зовут елизар, я студент 4 курса программной инженерии, группа ПИ-18-2

Тема разработка MVP системы машинного обучения для прогнозирования рентабельности кинобизнеса

**3 СЛАЙД:**

По данным отчета «Motion picture association» Американской ассоциации кинокомпаний, за 2019 год суммарные кассовые сборы всех вышедших в мировой кинотеатральный прокат фильмов составили 98,1 млрд $, что почти на 7 процентов больше, чем в 2018 году. Однако в 2020 году из-за сложившейся эпидемиологической ситуации во всем мире кассовые сборы резко упали и составили 80,8 млрд $, что на 18% меньше чем в 2019 году. Со снятием ограничений на посещение общественных мест в 2021 году, кассовые сборы возросли до почти 100 млрд $ но так как большинство зрителей перешло на потребление цифровых копий кинокартин, кассовые сборы театрального проката так и не наверстали упущенное

Таким образом цифровой прокат занимает наибольшую часть кассовых сборов кинопроката и доля такого проката лишь растет

К сожалению нельзя точно сказать на данный момент о ситуации в будущем, но явно присутствует тенденция на переход от театрального проката на цифровой для точной оценки ситуации требуется как минимум получить результаты этого и следующего года

**4 СЛАЙД:**

В связи с тотальным запретом на посещение общественных мест соответственно и упали продажи билетов, а так же количество выпушенных кинокартин, реузльтаты представлены в таблице

количество проданных билетов в 2020 году, по данным сайта «The Numbers» составило 223 млн., против 1,3 млрд в 2019 году и 496 млн. в 2021 году

Актуальность данной выпускной квалификационной работы заключается в том, что выпуск кинокартины в наши дни является большим риском, не только для киностудий и режиссеров, но и для инвесторов, система прогнозирования способна снизить риски неудачных инвестиций в кинопроизводство на ранних этапах создания кинокартины, а также спрогнозировать целесообразность выпуска кинокартины на экраны кинотеатров с сопутствующими затратами на рекламу и аренду кинозалов или ограничить выпуск лишь в цифровом формате.

**5 СЛАЙД:**

***Объектом*** исследования в данной работе являются алгоритмы систем машинного обучения.

***Предметом*** исследования выступает разработка алгоритма системы машинного обучения.

***Целью*** данной работы является разработка системы машинного обучения для прогнозирования рентабельности кинобизнеса.

Задачи для достижения поставленной цели:

Провести анализ систем машинного обучения и методов прогнозирования.

Провести анализ особенностей кинобизнеса.

Провести анализ особенностей построения систем машинного обучения.

Формализовать существующие методы прогнозирования рентабельности кинобизнеса.

Конкретизировать требования к разрабатываемой системе.

Спроектировать систему.

Спроектировать модуль сбора и систематизации данных.

Спроектировать системы машинного обучения.

Спроектировать графический интерфейс разрабатываемой системы.

Реализовать систему машинного обучения для прогнозирования рентабельности кинобизнеса.

Обосновать выбор технологий для разработки.

Реализовать модуль сбора и систематизации данных.

Реализовать систему машинного обучения.

Реализовать графический интерфейс разрабатываемой системы.

Провести тестирование системы.

**6 СЛАЙД:**

На слайде отображены основные этапы выпуска кинофильмов, помимо этих этапов так же стоит выделить сбор инвестиций, презентацию концепции и продвижения, за весь фильм от его начала до его конца отвечает продюсер.

В обязанности продюсера входит просчитывание: сколько кинотеатров купят фильм; сколько зрителей его посмотрят; насколько он окупится и какую принесет прибыль. Составление прогнозов относительно прибыли фильма является одной из самых важных частей создания фильмов, оцениваются риски и выявляются элементы на которые необходимо сфокусировать внимание.

Разрабатываемая система позволить на ранних этапах дать примерное подтверждение успешности фильма.

**7 СЛАЙД:**

По итогам сравнения видов нейронных сетей в ПРИЛОЖЕНИЕ B и главе 1.2.1 была составлена таблица, содержащая нейронные сети, способные решить задачу разрабатываемой системы, а именно задачу прогнозирования. Данные нейронные системы архитектурно являются разновидностями персептрона

Как итог была выбрана нейронная система вида персептрон.

**8 СЛАЙД:**

На основе анализа предыдущих работ и исследований выделены входные данные, представленные на слайде, которые будут рассмотрены и отсортированы во 2 главе данной работы для последующего обучения нейронной сети

**9 СЛАЙД:**

На момент написания аналитического обзора, данной дипломной работы, после сравнительного анализа предыдущих исследований в выбранной предметной области, было выявлено отсутствие общедоступных программных решений для прогнозирования кассовых сборов фильма. Единственным, найденным решением, является разработанное веб приложение, в рамках исследования схожей темы, а именно «Прогнозирование кассовых сборов фильма» в «Пермской научной школе искусственного интеллекта».

**10 СЛАЙД:**

**«Нейросимулятор 5.0»** – в программа-симулятор разработанная и зарегистрированная Ф.М. Черепановым и Л.Н. Ясницким в 2014 году. Программа предназначена для создания небольших нейронных сетей для решения любых задач, таким образом программа лишь косвенно подходит под выбранную предметную область

Преимущества данной программы состоят в функционале и гибкости системы, с возможностью подстраивать нейронную сеть программы под разные задачи. К минусам следует отнести сложность программы, как в виде интерфейса для рядового пользователя, так и в виде обучения самой нейронной системы.

Так же стоит учесть что данная программа полноценно не подходит под аналог разрабатываемой системы, но подходит для первичного обучения и тестирования входного множества и нейронной сети.

**11 СЛАЙД:**

Реализация программного продукта для прогнозирования рентабельности кинобизнеса будет проводится в несколько этапов, представленных на слайде

Для чего необходимо разработать следующие модули:

Модуль сбора данных обучающего множества (этап сбора данных).

Модуль очистки, систематизации и нормализации данных (этап подготовки и нормализации данных).

Модуль нейронной сети (этап обучения, этап проверки адекватности обучения, этап корректировки параметров).

Пользовательский модуль программного продукта (этап вербализации нейронной сети).

Требования к модулям описаны в главе 1.5 документа а также конкретизированы в техническом задании в приложении D

**12 СЛАЙД:**

На слайде представлены требования к входным и выходным данным

**13 СЛАЙД:**

Помимо технического задания была разработана диаграмма прецедентов в нотации UML

В диаграмме присутствует 2 актора: пользователь, который имеет доступ пользовательскому интерфейсу и может вводить данные и нейронная сеть, которая принимает введенные данные пользователем, обрабатывает их и выводит результат. В диаграмме также присутствует артефакт «Обученная выборка», как файл, с которым взаимодействует нейронная сеть для корректной работы внутри системы

Описание каждого прецедента приведено в приложении E

**14 СЛАЙД:**

Языком программирования был выбран Python из-за наличия доступных способов сбора данных и в основном построения нейронной сети так как на других языках разработка нейронной сети займет не один месяц и не меньше 1000 человеко-часов, для интерфейса так же был выбран Python для более удобного взаимодействия с нейронной сетью

Средой разработки является PyCharm так как компания разработчик JetBrains работает совместно с ниу вшэ

**15 СЛАЙД:**

Построенная нейронная сеть состоит из 3 слоев, в котором 1 входной слой с 11 нейронами, которые отображают входные параметры с активационной функцией вида сигмоида, 1 скрытый слой с 3 нейронами с сигмоидной активацией и 1 выходной слой для результата с линейной активацией. Для обучения было проведено 100 итераций. Результат обучения:

Среднее квадратичное отношение – 19%

Средняя относительная погрешность – 15.8%

Среднее квадратичное отношение теста – 24%

Средняя относительная погрешность теста – 19.8%

**16 СЛАЙД:**

Интерфейс системы будет аналогичен веб приложению представленному как аналог на одном из предыдущих слайдов

Интерфейс будет соответствовать требованиям MVP системы, а то есть иметь лишь набор базовых, необходимых для работы функций

**17 СЛАЙД:**

Так как интерфейс находится на этапе разработки то и его тестирование показано пока что быть не может

На данный момент нейронная сеть работает либо через консольное приложение PyCharm или через Google Collab, для запуска необходимо загрузить набор данных и переместить его в папку content