

Министерство образования Республики Беларусь  
Учреждение образования  
«Белорусский государственный университет  
информатики и радиоэлектроники»

УДК 004.3.049.77

ДОЛМАТОВИЧ  
Алина Сергеевна

**Программное средство для прогнозирования стоимости аренды  
недвижимости**

**АВТОРЕФЕРАТ**

диссертации на соискание степени  
магистра информатики и вычислительной техники

по специальности 1-40 81 04 – Обработка больших объемов информации

Научный руководитель

Калугина М.А.

канд.физ.-мат. наук., доцент

Минск 2020

Работа выполнена на кафедре информатики учреждения образования «Белорусский государственный университет информатики и радиоэлектроники»

Научный руководитель: **КАЛУГИНА Марина Алексеевна**,  
кандидат физико-математических наук, доцент,  
доцент кафедры информатики учреждения образования «Белорусский государственный университет информатики и радиоэлектроники»

Рецензент: **ЛАВРЕНОВ Александр Николаевич**  
кандидат физико-математических наук, доцент,  
доцент кафедры информационных технологий в образовании учреждения образования «Белорусский государственный педагогический университет»

Защита диссертации состоится «26» июня 2020 г. года в 15<sup>00</sup> часов на заседании Государственной экзаменационной комиссии по защите магистерских диссертаций в учреждении образования «Белорусский государственный университет информатики и радиоэлектроники» по адресу: 220013, Минск, ул. Гикало, 9, копр. 4, ауд. 111, тел. 293-85-91, e-mail: [inform@bsuir.by](mailto:inform@bsuir.by)

С диссертацией можно ознакомиться в библиотеке учреждения образования «Белорусский государственный университет информатики и радиоэлектроники».

## ВВЕДЕНИЕ

В современном мире особую важность приобрела проблема прогнозирования основных параметров рынка. Определение справедливой цены на аренду недвижимости является одной из основных проблем как для арендаторов, так и для арендодателей по той причине, что подобного рода задача является многофакторной. Помимо того, важным аспектом в жизни людей в настоящее время является их собственный мобильный телефон, который каждый человек всегда имеет при себе. Отсюда можно сделать предположение, что мобильное приложение может стать хорошим помощником для людей в вопросах определения стоимости арендного жилья.

Рыночная стоимость представляет собой наиболее вероятную цену, по которой объект сделки может быть отчужден на дату момента заключения сделки в условиях рыночной конкуренции.

На стоимость недвижимости оказывает влияние три уровня факторов: социальные, экономические, политические. Представленные факторы носят общий характер и не связаны с конкретным объектом недвижимости, однако они оказывают влияние на процессы, происходящие на рынке коммерческой недвижимости, на уровень спроса и предложения;

Местоположение объекта, физические характеристики объекта, временные факторы, условия продажи, аренды и финансирования. Данные факторы носят локальный характер и связаны с особенностями конкретной территории и района;

Финансово-эксплуатационные и архитектурно-строительные факторы, которые непосредственно связаны с объектом недвижимости и его техническими характеристиками.

Для получения достоверных результатов в процессе оценки и прогнозирования рыночной стоимости аренды коммерческой недвижимости необходимо тщательно выявлять, анализировать и учитывать все представленные уровни факторов.

В качестве основного инструмента исследования в этом случае выступает задача восстановления линейной регрессии как отдельный вариант из множества алгоритмов машинного обучения. Линейная регрессия является одним из базовых алгоритмов для многих областей, связанных с анализом данных. Причина тому – простота алгоритма, которая способствует широкому применению уже многие десятки, если не сотни, лет. Основная идея заключается в восстановлении линейной зависимости одной переменной от набора других переменных. Модели на основе линейной регрессии позволяют определять максимально близкую к реальной цену на недвижимость.

На сегодняшний день существует небольшое число работ в области предсказания цен на недвижимость. Наиболее интересными оказались работы Sumit C., Trivikraman T., и др. «Определение цены на

недвижимость», Robert F., David M. и др. «Подход к оценке», Weikun Zhao, Cao Sun и др. «Построение модели». Также были рассмотрены онлайн-сервисы Zillow и Flatorial.

## **ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ**

### **Цель и задачи исследования**

*Целью* диссертационной работы является разработка мобильного программного средства на платформе iOS для решения задачи регрессии в контексте предсказания стоимости аренды недвижимости.

Для достижения поставленной цели необходимо решить следующие задачи:

- 1) изучить существующие аналоги, ресурсы, алгоритмы;
- 2) найти достаточную для работы базу данных;
- 3) проанализировать и выбрать необходимые инструменты для реализации программного средства;
- 4) обработать полученные данные;
- 5) построить регрессионную модель;
- 6) проверить модель на реальных данных;
- 7) провести измерение полученных результатов;
- 8) разработать архитектуру программного средства;
- 9) реализовать программное средство для предсказания стоимости аренды недвижимости;
- 10) провести экспериментальные исследования разработанной системы.

*Объектом* исследования является прогнозирование стоимости аренды недвижимости.

*Предметом* исследования является программное средство для прогнозирования стоимости аренды недвижимости.

Основной *гипотезой*, положенной в основу диссертационной работы, является возможность использования мобильного устройства с ОС iOS для решения задачи ввода, обработки и анализа информации об объекте недвижимости с целью предсказания стоимости ее аренды.

### **Степень разработанности проблемы**

Основные исследования данных вопросов были найдены у Sumit C., Trivikraman T., Robert F., David M., Weikun Zhao, Cao Sun и др.

В ходе анализа этих исследований и ресурсов по предсказанию цен на недвижимость были обнаружены следующие особенности:

1. Отсутствует универсальный подход для решения задачи.
2. Результаты сильно зависят от качества базы данных.

3. Данные брались из открытых источников, по этой причине алгоритмы имели ограничения.
4. Комбинации различных моделей часто улучшают результат.

Стоит заметить, что не было найдено ни одного мобильного приложения, которое способно каким-то образом помочь в оценке стоимости аренды недвижимости.

### **Область исследования**

Содержание диссертации соответствует образовательному стандарту высшего образования второй ступени (магистратуры) ОСВО 1-40 81 04-2020 специальности 1-40 81 04 – “Обработка больших объемов информации”.

### **Теоретическая и методологическая основа исследования**

В основу диссертации легли работы зарубежных ученых в области решения задачи восстановления регрессии, а также анализ онлайн-сервисов, технической документации по рассматриваемой тематике.

*Информационная база* исследования сформирована на основе литературы, открытой информации, сведений из электронных ресурсов, а также материалов научных конференций и семинаров.

### **Научная новизна**

*Научная новизна* и значимость полученных результатов работы заключается в разработке мобильного приложения, способного предсказывать стоимость для аренды недвижимости путем ввода значений для ряда параметров, характеризующих интересующий объект.

*Теоретическая значимость* работы заключается в детальном анализе алгоритмов, способов оценки алгоритмов машинного обучения и взаимодействия мобильного приложения с моделями машинного обучения.

*Практическая значимость* диссертации заключается в разработанном мобильном приложении для оценки стоимости аренды недвижимости.

### **Публикации**

По теме диссертации опубликованы 3 печатные работы, из них 2 работы на сайте компании VironIT, 1 работа в сборнике тезисов докладов 56-ой научной конференции аспирантов, магистрантов и студентов БГУИР.

### **Структура и объем работы**

Диссертация состоит из введения, общей характеристики работы, четырех глав, заключения, списка использованных источников, списка публикаций автора и трех приложений.

В первой главе представлен анализ предметной области и обзор используемой базы данных. Рассматриваются существующие аналоги с их краткими характеристиками.

Вторая глава дает краткую характеристику используемых технологий и описывает причины, по которым они были выбраны.

В третьей главе приведены сведения о ходе выполнения работы и важнейших особенностях применений используемых технологий.

В четвертой главе описана реализация мобильного средства с демонстрацией работы.

Общий объем работы составляет 96 страниц, из которых основного текста — 60 страниц, 20 рисунков на 19 страницах, список использованных источников из 30 наименований на 2 страницах и 3 приложения на 36 страницах.

## ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** рассмотрено современное состояние проблемы предсказания цен на аренду недвижимости, указаны основные направления исследований, проводимых по данной тематике, а также описано обоснование актуальности темы.

В **общей характеристике работы** показана актуальность проводимых исследований, степень разработанности проблемы, сформулированы цель и задачи диссертации, обозначена область исследований, научная (теоретическая и практическая) значимость исследований.

В **первой главе** приведен анализ предметной области. Прежде, чем приступить к разработке собственного решения, были найдены и изучены существующие аналоги. Рассмотрены наиболее интересные работы.

В работе Sumit S., Trivikraman T., и др. «Определение цены на недвижимость» был предложен новый подход к регрессии, в котором предсказываемые переменные также зависят от множества других переменных. Был представлен подход, названный LME, который комбинирует обучаемую параметрическую модель и модель непараметрического многообразия, чтобы сделать прогноз.

В работе Robert F., David M. и др. «Подход к оценке» представлена и оценена модель рынка вторичного жилья. Данные представляют собой месячные временные ряды для пригорода Сан-Диего. Они оцениваются по максимальной вероятности с использованием EM-алгоритма, основанного на фильтрации и сглаживании Калмана.

В работе Weikun Zhao, Cao Sun и др. «Построение модели» речь идет о том, что рынок подержанного жилья является основополагающим рынка недвижимости, поскольку покупатели подержанных домов обычно являются теми, кто действительно хочет там жить, тогда как финансовые инвестиции и спекуляция не являются их целями. Таким образом, цена подержанных домов отражают реальный спрос на рынке жилья. В этой статье алгоритмы,

связанные с KNN, применяются для изучения проблем, связанных с ценой на подержанный дом.

Также были рассмотрены онлайн-сервисы Zillow и Flatorial. Это сервисы, автоматически оценивающие стоимость квартир на основе методов машинного обучения.

**Во второй главе** дана краткая характеристика используемыми технологиям с обоснованием выбора. Для разработки модели машинного обучения было принято решение использовать Python. По сути, машинное обучение — это технология, которая помогает приложениям на основе искусственного интеллекта обучаться и выдавать результаты автоматически, без человеческого вмешательства. Python отличается отличной производительностью при обработке данных.

Помимо простоты, у Python есть еще один плюс — он довольно легко взаимодействует с другими языками. Этот факт тоже стоит учитывать по той причине, что необходимо связывать модель машинного обучения с кодом для мобильного приложения для iOS.

Для создания модели машинного обучения были рассмотрены такие технологии, как Tensorflow, PyTorch и Keras. Для каждой из них были изложены свои достоинства и недостатки, на основании которых в рамках данной работы было принято решение использовать Keras.

Для разработки мобильного приложения выбран Swift. CoreML — для интеграции модели машинного обучения.

Swift для написания приложения требует меньше кода, хотя бы просто потому, что здесь реализован упрощенный принцип работы с повторяющимися строками и заявлениями; он удобен для чтения (стандартное достоинство любого современного языка); имеет больше возможностей по сравнению с Objective-C, в частности возможность управлять памятью; отличается повышенной безопасностью (это выражается и в обработке указателей, и в дотошности компилятора, и в том, что в саму компиляцию можно встроить опциональную переменную nil для обеспечения обратной связи).

**В третьей главе** рассмотрены алгоритмы для решения задачи восстановления линейной регрессии, измерены полученные результаты. Регрессионная модель показала лучший результат на ансамбле алгоритмов Adadelta и MSE с коэффициентом смешанной корреляции 0.81, что является достаточно высоким показателем оптимальности. На основании вычислений точности рассмотренных моделей можно сделать вывод, что этот ансамбль наиболее точно позволяет определять цену на аренду недвижимости.

Измерения качества алгоритмов проводились с использованием кроссвалидации. В качестве метрик качества работы алгоритма использовались MSE, MAE и MAPE.

Для того, чтобы можно было использовать полученную модель в iOS приложении, необходимо было конвертировать полученную Keras-модель в модель CoreML.

В **четвертой главе** рассматривается разработка мобильного приложения, начиная с ее архитектуры. Рассматривается архитектурный паттерн MVC, рекомендуемый Apple для использования в своих приложениях. Само мобильное приложение представляет собой простое оконное приложение со списком параметров, которые необходимо ввести пользователю об интересующем его объекте недвижимости. Пользовательский интерфейс представлен на рисунке 2.

## **ЗАКЛЮЧЕНИЕ**

### **Основные научные результаты диссертации**

В ходе диссертационного исследования был проведен анализ предметной области, в том числе были проанализированы существующие аналоги.

На основе анализа предметной области и найденной в открытом доступе базы данных определены оптимальные языки и технологии для разработки, такие как Python, Swift, Keras, Core ML.

С использованием выбранных технологий разработан алгоритм определения цены на недвижимость на основе регрессионной модели с достаточно высокими показателями точности и проверен на реальных данных.

Было спроектировано программное средство, разработана программная архитектура, алгоритмы основных функций программного средства, в том числе общий алгоритм работы приложения.

Итогом выполнения магистерской диссертации является мобильное программное средство для оценки стоимости аренды недвижимости.



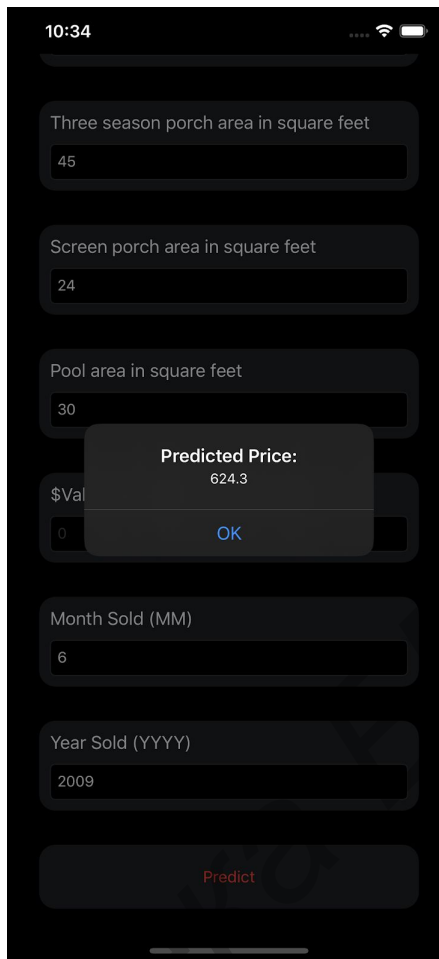


Рисунок 2 - Пользовательский интерфейс приложения

### СПИСОК ПУБЛИКАЦИЙ СОИСКАТЕЛЯ

1-А. Долматович А.С. Мобильное приложение для контроля питания / А.С. Долматович // Компьютерные системы и сети: 56-я научная конференция аспирантов, магистрантов и студентов, Минск, Республика Беларусь, 21-24 апреля 2020 г. / Белорусский государственный университет информатики и радиоэлектроники. – Минск, 2020 – С.134-135.

2-А. Долматович А.С. Functional Programming in Swift / А.С. Долматович // VironIT - 2019 - URL: <https://vironit.com/functional-programming-in-swift/>

3-А. Долматович А.С. Machine Learning in iOS/ А.С. Долматович // VironIT - 2018 - URL: <https://vironit.com/machine-learning-in-ios/>