МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное образовательное учреждение

высшего образования

**«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»**

**(ФГБОУ ВО «КубГУ»)**

**Факультет компьютерных технологий и прикладной математики**

**Кафедра прикладной математики**

**КУРСОВАЯ РАБОТА**

**МАШИННОЕ ОБУЧЕНИЕ ДЛЯ ОЦЕНКИ СТОИМОСТИ КВАРТИР В КРАСНОДАРЕ**

Работу выполнил \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ П.Ю. Захаров

(подпись)

Направление подготовки 09.03.03 Прикладная информатика

(код, наименование)

Направленность Прикладная информатика в экономике

(наименование)

Научный руководитель

канд. физ.-мат. наук, доц. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ А.Д. Колотий

(подпись)

Нормоконтролер

канд. физ.–мат. наук \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Г.В. Калайдина

(подпись)

Краснодар

2021

**РЕФЕРАТ**

Курсовая работа 24 с., 15 рис., 2 табл., 6 источников.

МАШИННОЕ ОБУЧЕНИЕ, ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ, ИНС, ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ, ПАРСИНГ, СТАТИСТИЧЕСКИЙ АНАЛИЗ, ЗАДАЧА РЕГРЕССИИ

В данной работе исследуется рынок вторичного жилья в г. Краснодар с целью разработки математической модели оценки стоимости квартир на основе искусственной нейронной сети.

Перед обучением нейронной сети для предсказания стоимости квартир затрагиваются задачи сбора информации об объявлениях продажи из внешних источников и проводится анализ факторов, влияющих на цену жилья.

В результате происходит обучение нейронной сети на полученных и обработанных данных, оценка потерь обученной модели. Напоследок приведено несколько примеров оценки стоимости квартир.

**СОДЕРЖАНИЕ**

[Введение 4](#_Toc66543732)

[1 Постановка задачи 5](#_Toc66543733)

[2 Нейронные сети 6](#_Toc66543734)

[2.1 История машинного обучения и ИНС 6](#_Toc66543735)

[2.2 Искусственный нейрон 7](#_Toc66543736)

[2.2.1 Перцептрон 10](#_Toc66543737)

[2.2.2 Функции активации 12](#_Toc66543738)

[3 Задача классификации 15](#_Toc66543739)

[3.1 Разбор поставленной задачи 15](#_Toc66543740)

[3.2 Метод ближайших соседей 16](#_Toc66543741)

[Заключение 22](#_Toc66543742)

[Список использованных источников 23](#_Toc66543743)

[Приложение А 24](#_Toc66543744)

**ВВЕДЕНИЕ**

Вторичный рынок недвижимости — это совокупность участников и договоров между ними по покупке, продаже, аренде и т. д. объектов вторичной недвижимости. Дом или квартира являются вторичным жильём, если право собственности на них оформляется более одного раза.

Вторичное жилье пользуется большим спросом в России. По данным Росреестра, в декабре 2020 года впервые превышен показатель в 20000 продаж вторичного жилья за месяц в Москве [1]. Количество сделок на вторичном рынке превысило количество сделок на первичном за этот же период в Москве, и это при общем росте спроса на недвижимость за последний год [2].

Это актуально и для Краснодара. По данным Росриэлт, количество объявлений о продаже квартир на вторичном рынке почти в 10 раз превышает количество объявлений для новостроек [3]. Здесь же сделаем замечание, что средняя цена за квадратный метр в Краснодаре выросла примерно на 20% за год.

Одним из самых важных критериев при выборе квартиры является её стоимость. Как для продавца, так и для покупателя важно знать рыночную стоимость жилья, т. е. усреднённую стоимость квартир с похожими параметрами. Обычно этим занимаются оценщики и риэлторы. Но нужно понимать, что даже в таком случае цена может значительно отличаться от рыночной. Например, она будет ниже рыночной если стоит цель продать квартиру как можно скорее.

**1 Постановка задачи**

В данной курсовой работе поставлены следующие задачи:

* обзор и ознакомление с историей машинного обучения и ИНС;
* изучение модели нейронных сетей и теории обучения нейронных сетей;
* изучение проблемы задачи классификации;
* метод ближайших соседей (kNN);
* пример решения задачи классификации при помощи метода kNN;
* изучение вопроса об изменении поставленной задачи в зависимости от рода входных данных;
* рассмотрение принципов распознавания объектов на изображениях;

**2 Парсинг сайта**

**2.1 История машинного обучения и ИНС**

Исследования в области искусственного интеллекта пережили

**ЗАКЛЮЧЕНИЕ**

Была рассмотрена лишь малая часть задач, решаемых при помощи нейронных сетей. Даже задача классификации не всегда решается методом kNN. Например, для классификации изображений используют другой подход – свёрточные нейронные сети (CNN). Таким образом, можно сделать вывод, что задача классификации очень чувствительна к роду входных данных.

**СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ**

1. Статья Росреестра о рекорде продаж на вторичном рынке жилья за месяц. URL: – https://rosreestr.gov.ru/site/press/news/v-moskve-vpervye-prevyshen-pokazatel-v-20-tys-registratsiy-na-vtorichnom-rynke-zhilya-/ (дата обращения 24.05.2021)
2. Статья РБК о рекордном спросе на новостройки в Москве. URL: – https://realty.rbc.ru/news/60092b869a7947e5c4a109bb (дата обращения 24.05.2021)
3. Росриэлт. Цены на недвижимость в Краснодаре. URL: – https://rosrealt.ru/krasnodar/cena (дата обращения 24.05.2021)
4. Шолле Ф. Глубокое обучение на Python. СПб.: Питер, 2020. – 400 с.
5. Вьюгин В. В. Математические основы машинного обучения и прогнозирования. М.: МЦНМО, 2013. – 304 с.
6. Samsung Research Russia. Онлайн–курс «Нейронные сети и компьютерное зрение». URL: – https://stepik.org/course/50352/info (дата обращения 27.02.2021)
7. Типология задач обучения по прецедентам. URL: – http://www.machinelearning.ru/wiki/index.php?title=Машинное\_обучение (дата обращения 27.02.2021)
8. Deep Learning School. Онлайн–курс «Глубокое обучение». URL: – https://stepik.org/course/65388/info (дата обращения 27.02.2021)
9. Датасет лесного покрытия / Covertype Data Set. URL: – http://archive.ics.uci.edu/ml/datasets/Covertype (дата обращения 27.02.2021)

# ПРИЛОЖЕНИЕ А

**Код программы**

import pandas as pd

import numpy as np

from sklearn.model\_selection import train\_test\_split

Рисунок А.1 – Полный код программы