

**Министерство науки и высшего образования Российской Федерации  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ  
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО  
ITMO University**

**ЗАДАНИЕ НА ВЫПУСКНУЮ КВАЛИФИКАЦИОННУЮ РАБОТУ /  
OBJECTIVES FOR A GRADUATION THESIS**

**Обучающийся / Student** Колесникова Мария Алексеевна  
**Факультет/институт/кластер/ Faculty/Institute/Cluster** факультет прикладной информатики  
**Группа/Group** K34212  
**Направление подготовки/ Subject area** 11.03.02 Инфокоммуникационные технологии и системы связи  
**Образовательная программа / Educational program** Программирование в инфокоммуникационных системах 2021  
**Язык реализации ОП / Language of the educational program** Русский  
**Квалификация/ Degree level** Бакалавр  
**Тема ВКР/ Thesis topic** Разработка классификатора компаний на основе анализа ассоциированных открытых проектов  
**Руководитель ВКР/ Thesis supervisor** Капитонов Александр Александрович, доцент, кандидат технических наук, Университет ИТМО, факультет прикладной информатики, доцент (квалификационная категория "ординарный доцент")

**Характеристика темы ВКР / Description of thesis subject (topic)**

**Тема в области фундаментальных исследований / Subject of fundamental research:** да / yes

**Тема в области прикладных исследований / Subject of applied research:** нет / not

**Основные вопросы, подлежащие разработке / Key issues to be analyzed**

Техническое задание: Разработать программное решение для автоматизированной классификации компаний на основе анализа их ассоциированных открытых проектов с использованием методов машинного обучения и анализа данных.

Исходные данные к работе: Данные из открытых репозиторий GitHub. Алгоритмы машинного обучения: логистическая регрессия, метод опорных векторов, случайный лес, градиентный бустинг, k-ближайших соседей, нейронные сети и др. Метрики оценки качества модели классификации: Accuracy, Precision, Recall, F1-score.

Содержание работы: Анализ предметной области и формирование требований к классификатору. Разработка методики классификации компаний. Реализация программного решения. Оценка и анализ результатов работы модели классификатора. Подготовка рекомендаций по улучшению модели.

Цель: Создание классификатора компаний на основе анализа их ассоциированных открытых проектов.

Задачи работы: Разработать программное решения для автоматизированного сбора данных с платформы Github с помощью Github API. Исследовать закономерности и определить

признаки, по которым будет осуществляться классификация компаний. Исследовать алгоритмы машинного обучения, применимые к задаче классификации, и выбрать оптимальный. Разработать программный модуль для автоматизированной классификации компаний. Провести тестирование модели на выборке данных и оценить качество классификации. Подготовить рекомендации по улучшению модели и возможному расширению функциональности.

Перечень подлежащих разработке вопросов: Разработка программного решения для сбора данных с применением Github API. Подготовка данных (очистка, нормализация, заполнение пропусков). Выбор алгоритма машинного обучения и построение модели классификации. Разработка классификационной модели с использованием Python и библиотек для анализа данных (pandas, scikit-learn и др.). Тестирование и оценка эффективности модели. Анализ полученных результатов. Разработка рекомендаций по улучшению модели.

Рекомендуемые материалы и пособия для выполнения работы: Документация API GitHub. Документация по библиотекам pandas, scikit-learn и другим инструментам машинного обучения.

### **Форма представления материалов ВКР / Format(s) of thesis materials:**

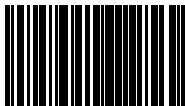
Отчет, презентация

**Дата выдачи задания / Assignment issued on:** 21.02.2025

**Срок представления готовой ВКР / Deadline for final edition of the thesis** 24.05.2025

### **СОГЛАСОВАНО / AGREED:**

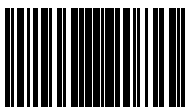
Руководитель ВКР/  
Thesis supervisor

Документ подписан	
Капитонов Александр Александрович	
13.03.2025	

(эл. подпись)

Капитонов  
Александр  
Александрович

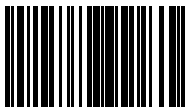
Задание принял к  
исполнению/ Objectives  
assumed BY

Документ подписан	
Колесникова Мария Алексеевна	
05.05.2025	

(эл. подпись)

Колесникова  
Мария  
Алексеевна

Руководитель ОП/ Head  
of educational program

Документ подписан	
Береснев Артем Дмитриевич	
07.05.2025	

(эл. подпись)

Береснев Артем  
Дмитриевич