



МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение  
высшего образования

«МИРЭА - Российский технологический университет»

РТУ МИРЭА

Институт искусственного интеллекта

Кафедра высшей математики

**ОТЧЁТ ПО НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ**  
**(получение первичных навыков научно-исследовательской работы)**

**Тема НИР:** Предсказание стоимости медицинской страховки (kaggle.com)  
приказ университета о направлении на практику  
от «9» февраля 2022 г. № 1038 - С

Отчет представлен к  
рассмотрению:

Студент группы КМБО-01-  
21

Сиразетдинов Р.Д.  
(расшифровка подписи)  
«\_\_» \_\_\_\_\_ 2022 г.

Отчет утвержден.  
Допущен к защите:

Руководитель практики от  
кафедры

Петрусеви́ч Д.А.  
(расшифровка подписи)  
«\_\_» \_\_\_\_\_ 2022 г.

Москва 2022



МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение  
высшего образования

«МИРЭА - Российский технологический университет»

РТУ МИРЭА

**ЗАДАНИЕ**

**на НАУЧНО-ИССЛЕДОВАТЕЛЬСКУЮ РАБОТУ**

**(получение первичных навыков научно-исследовательской работы)**

**Студенту 1 курса учебной группы КМБО-01-21 института искусственного интеллекта  
Сиразетдинову Рустему Дамировичу**

(фамилия, имя и отчество)

**Место и время НИР:** Институт искусственного интеллекта, кафедра высшей математики

**Время НИР:** с «09» февраля 2022 по «31» мая 2022

**Должность на НИР:** практикант

**1. ЦЕЛЕВАЯ УСТАНОВКА:** изучение основ анализа данных и машинного обучения

**2. СОДЕРЖАНИЕ НИР:**

2.1 Изучить: литературу и практические примеры по темам: 1) построение линейной регрессии, 2) использование метода главных компонент, 3) поиск и устранение линейной зависимости в данных, 4) основы нормализации данных, 5) методы классификации и кластеризации («решающее дерево», «случайный лес», «k ближайших соседей»).

2.2 Практически выполнить: 1) снижение размерности исходных задач при помощи метода главных компонент при возможности; построение линейной регрессии для некоторого параметра, исключение регрессоров, не коррелирующих с объясняемой переменной; решение задачи классификации или кластеризации на основе открытого набора данных с ресурса kaggle.com

2.3 Ознакомиться: с применением метода главных компонент; методов классификации («решающего дерева», «случайного леса»); методов кластеризации («k ближайших соседей»); построением модели линейной регрессии

**3. ДОПОЛНИТЕЛЬНОЕ ЗАДАНИЕ:** предсказание стоимости медицинской страховки (kaggle.com)

**4. ОГРАНИЗАЦИОННО-МЕТОДИЧЕСКИЕ УКАЗАНИЯ:** построить модель предсказания, оценить вклад каждой компоненты. Есть ли выбросы среди данных, какие? Применить алгоритмы кластеризации. Что общего между объектами в каждом кластере?

Заведующий кафедрой

высшей математики

«09» февраля 2022 г.

Ю.И.Худак

СОГЛАСОВАНО

Руководитель практики от кафедры:

«09» февраля 2022 г.

(подпись)

(Петрусеви́ч Д.А.)

(фамилия и инициалы)

Задание получил:

«09» февраля 2022 г.

(подпись)

(Сиразетдинов Р.Д.)

(фамилия и инициалы)

**ИНСТРУКТАЖ ПРОВЕДЕН:**

| <b>Вид мероприятия</b>            | <b>ФИО ответственного,<br/>подпись, дата</b>        | <b>ФИО студента, подпись,<br/>дата</b>               |
|-----------------------------------|---|--|
| Охрана труда                      | <u>Петрусеви́ч Д.А.</u><br><br>«09» февраля 2022 г. | <u>Сиразетдинов Р.Д.</u><br><br>«09» февраля 2022 г. |
| Техника безопасности              | <u>Петрусеви́ч Д.А.</u><br><br>«09» февраля 2022 г. | <u>Сиразетдинов Р.Д.</u><br><br>«09» февраля 2022 г. |
| Пожарная безопасность             | <u>Петрусеви́ч Д.А.</u><br><br>«09» февраля 2022 г. | <u>Сиразетдинов Р.Д.</u><br><br>«09» февраля 2022 г. |
| Правила внутреннего<br>распорядка | <u>Петрусеви́ч Д.А.</u><br><br>«09» февраля 2022 г. | <u>Сиразетдинов Р.Д.</u><br><br>«09» февраля 2022 г. |



**МИНОБРНАУКИ РОССИИ**

Федеральное государственное бюджетное образовательное учреждение  
высшего образования

**«МИРЭА - Российский технологический университет»**

**РТУ МИРЭА**

**РАБОЧИЙ ГРАФИК ПРОВЕДЕНИЯ  
НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЫ**

**(получение первичных навыков научно-исследовательской работы)**

студента Сиразетдинова Р.Д. 1 курса группы КМБО-01-21 очной формы обучения,  
обучающегося по направлению подготовки 01.03.02 «Прикладная математика и  
информатика»,  
профиль «Математическое моделирование и вычислительная математика»

| Неделя | Сроки<br>выполнения | Этап  | Отметка о<br>выполнении |
|--------|---------------------|---|-------------------------|
| 1      | 09.02.2022          | Выбор темы НИР. Пройти инструктаж по технике безопасности                                 |                         |
| 1      | 09.02.2022          | Вводная установочная лекция   |                         |
| 2      | 14.02.2022          | Построение и оценка парной регрессии с помощью языка R                                    |                         |
| 3      | 21.02.2022          | Построение и оценка множественной регрессии с помощью языка R                             |                         |
| 4      | 28.02.2022          | Построение доверительных интервалов. Обработка факторных переменных. Мультиколлинеарность |                         |
| 5      | 07.03.2022          | Гетероскедастичность  |                         |
| 6      | 14.03.2022          | Классификация   |                         |
| 7      | 21.03.2022          | Кластеризация. Предобработка данных   |                         |
| 8      | 28.03.2022          | Метод главных компонент   |                         |
| 9      | 04.04.2022          | Ансамбли классификаторов.   |                         |

|    |            |  |  |
|----|------------|--|--|
|    |            | Беггинг. Бустинг   |  |
| 16 | 29.05.2022 | Представление отчётных материалов по НИР и их защита. Передача обобщённых материалов на кафедру для архивного хранения |  |
|    |            | Зачётная аттестация  |  |

**Согласовано:**

Заведующий кафедрой \_\_\_\_\_ / ФИО / Худак Ю.И.

Руководитель практики от кафедры \_\_\_\_\_ / ФИО / Петрусевич Д.А.

Обучающийся \_\_\_\_\_ / ФИО / Сиразетдинов Р.Д.