



**Федеральное государственное автономное образовательное
учреждение высшего образования
"Национальный исследовательский
университет "Высшая школа экономики"**

Факультет компьютерных наук

**Рабочая программа дисциплины
НИС «Прикладной анализ данных»**

для образовательной программы «Финансовые технологии и анализ
данных» направления 01.04.02
уровень магистр

Разработчик программы:
Масютин Алексей Александрович
amasyutin@hse.ru

Одобрена на заседании
Департамента больших данных и информационного поиска
Рук. департамента Подольский В.В. _____ « ____ » _____ 2018 г.

Утверждено Академическим советом
магистерской программы «Финансовые технологии и анализ данных»
« ____ » _____ 201_ г., № протокола _____

Академический руководитель
Масютин А.А. _____ « ____ » _____ 2018 г.

Настоящая программа не может быть использована другими подразделениями университета и другими вузами без разрешения подразделения-разработчика программы.



1. Область применения

Настоящая программа учебной дисциплины устанавливает минимальные требования к знаниям и умениям студента и определяет содержание и виды учебных занятий и отчетности.

Программа предназначена для преподавателей, ведущих данную дисциплину, учебных ассистентов и студентов направления подготовки 01.04.02 «Прикладная математика и информатика», обучающихся по образовательной программе «Финансовые технологии и анализ данных».

2. Цели освоения дисциплины

Дисциплина преподается с целью ознакомления слушателей с подходами к решению бизнес-задач в крупных финансовых организациях с применением методов математического моделирования и анализа данных, приобретения навыков постановки прикладных исследовательских задач, разработки новых подходов и методов анализа данных, а также презентации реализованных решений перед заказчиками (в форме курсовой или выпускной квалификационной работы).

В течение 1-ого года обучения в проведении дисциплины участвуют сотрудники подразделений Сбербанка и консалтинговых компаний с успешным опытом решения бизнес-задач на основе математического моделирования и анализа данных. В течение 2-ого года обучения слушатели изучают последние статьи по разработке новых методов машинного обучения, получают информацию о методах машинного обучения с подкреплением и выступают по тематике своих выпускных квалификационных работ.

Форма итогового контроля: в течение 1-ого года обучения слушатели выполняют:

- 2 домашних проекта;
- готовят презентацию по результатам, полученным в рамках подготовки своей курсовой работы.

В течение 2-ого года обучения слушатели:

- проводят презентацию выбранной научной статьи, посвященной методам анализа данных, машинному обучению или финансовым технологиям, опубликованной не ранее 2 лет до проведения дисциплины;
- проводят презентацию результатов, полученных в рамках подготовки ВКР;
- выполняют домашнее задание по методам машинного обучения с подкреплением.

Итоговая оценка рассчитывается по следующей формуле:

$$O_{\text{итог}} = 0,4 * O_{1\text{-ый год}} + 0,6 * O_{2\text{-ой год}}$$

где $O_{1\text{-ый (2-ой) год}}$ – оценка за 1-ый (2-ой) год обучения.

$$O_{1\text{-ый год}} = 0,1 * O_{1\text{-ый проект}} + 0,8 * O_{2\text{-ой проект}} + 0,1 * O_{\text{презентация}}$$

$$O_{2\text{-ой год}} = 0,4 * O_{\text{ДЗ}} + 0,2 * O_{\text{през.статьи}} + 0,2 * O_{\text{през.ВКР}}$$

1-ый домашний проект является индивидуальным заданием по теме анализа временных данных, оценивается корректность примененного метода и точность полученного решения. 2-ой домашний проект является групповым заданием, оценивается корректность примененных методов, бизнес-логика предлагаемого решения и презентация решения. Презентация по курсовой работе оценивается по критериям структурированности и ясности сообщения, которое позволяет понять читателю примененные методы и эксперименты с данными, проведенные для оценки эффективности алгоритмов анализа данных, а также по способам визуализации, которые использует слушатель.

Домашнее задание оценивается по критериям корректности предложенного решения и проверяет, насколько слушатель освоил методы машинного обучения с подкреплением. Презентация научной статьи оценивается по критерию новизны, прикладному значению изученной работы для анализа данных в крупных финансовых организациях. Презентация по ВКР оценивается аналогично: по критериям структурированности и ясности сообщения, которое позволяет понять читателю примененные методы и эксперименты с данными, проведенные для оценки эффективности алгоритмов анализа данных, а также по способам визуализации, которые использует слушатель.

По окончании семинара слушатель будет:

- Знать особенности использования математического моделирования в крупных финансовых организациях;
- Уметь формулировать бизнес-задачу и переводить ее в термины задачи построения предиктивной модели;
- Уметь самостоятельно изучать и искать литературу, работы по проблематике предметной области,
- Уметь аргументировать выбор того или иного способа решения поставленной задачи,
- Уметь выступать с научно-исследовательскими докладами и презентациями практических бизнес-решений.

Список тем семинара:

- Методы машинного обучения для верификации прогнозов финансовых данных компаний в рамках CF-моделирования;
- Основные типы банковского кибермошенничества и степень последствий;
- Методы выявления и борьбы с кибермошенничеством на основе анализа транзакционных данных в онлайн канале;
- Продвинутое техники валидации моделей выявления кибермошенничества, методы локальной интерпретации и проблемы несбалансированности классов.