

**Федеральное государственное автономное образовательное
учреждение высшего образования
«Московский физико-технический институт
(национальный исследовательский университет)»**

УТВЕРЖДЕНО

**Директор физтех-школы
прикладной математики и
информатики**

А.М. Райгородский

	Рабочая программа дисциплины (модуля)
по дисциплине:	Финансовые технологии
по направлению:	Прикладная математика и информатика
профиль подготовки:	Комбинаторика и цифровая экономика центр дополнительного, дополнительного профессионального и онлайн-образования "Пуск" кафедра дискретной математики
курс:	2
квалификация:	магистр

Семестр, формы промежуточной аттестации: 4 (весенний) - Экзамен

Аудиторных часов: 60 всего, в том числе:

лекции: 30 час.

семинары: 30 час.

лабораторные занятия: 0 час.

Самостоятельная работа: 45 час.

Подготовка к экзамену: 30 час.

Всего часов: 135, всего зач. ед.: 3

Программу составил: А.В. Куликов, д-р физ.-мат. наук

Программа обсуждена на заседании кафедры дискретной математики 30.01.2025

Аннотация

Данный курс посвящен рассмотрению современных математических / статистических методов и методов машинного обучения и NLP и их применения к финансовым задачам. Основными задачами являются: различные модели анализа и предсказания временных рядов, применение NLP в финансах, нахождение оптимальных портфелей и оценка рисков. В курсе рассматриваются следующие вопросы:

введение базовых объектов теории финансов; введение мер риска и их использование для решения различных задач теории финансов; рассмотрение базовых объектов финансовой математики (фундаментальной и рыночной цены финансовых активов, первичных финансовых инструментов (акций и облигаций), а также производных финансовых инструментов (форвардов, фьючерсов, свопов, различных видов опционов);

нахождение цен различных финансовых инструментов, используя теорию арбитража в общей модели, колл-пут паритет, цены имеющихся на рынке опционов и т.д.;

описание и нахождение справедливых цен различных производных ценных бумаг для моделей, часто используемых в финансовой математике.

В качестве задач также будут рассмотрены вопросы, часто задаваемые на собеседованиях в финансовых организациях, а также методы, используемые для оценки риска и нахождения цен финансовых инструментов.

Вероятностные методы имеют широчайшее применение в этой области, поэтому в курсе также будут введены и рассмотрены важные элементы теории мартигалов и выпуклого анализа, а именно, условные математические ожидания, теория мартигалов.

1. Цели и задачи

Цель дисциплины

«Финансовые технологии» направлена на обучение основам современных математических / статистических методов и методов машинного обучения и NLP, которые имеют широчайшее применение в этой области.

Задачи дисциплины

научиться оперировать с базовыми объектами, применимыми при построении финансовых моделей;

заложить основы применения изученных методов в финансах
приобрести навык полного построения пайплайна модели.

2. Перечень формируемых компетенций

Освоение дисциплины направлено на формирование следующих компетенций:

Код и наименование компетенции	Индикаторы достижения компетенции
УК-3 Способен организовывать и руководить работой команды, вырабатывая командную стратегию для достижения поставленной цели	УК-3.1 Организует и координирует работу участников проекта, способствует конструктивному преодолению возникающих разногласий и конфликтов
ОПК-2 Способен совершенствовать и реализовывать новые математические методы решения прикладных задач	ОПК-2.1 Имеет представление о современном состоянии математических исследований в рамках тематической области своей профессиональной деятельности
ОПК-3 Способен разрабатывать математические модели и проводить их анализ при решении задач в области профессиональной деятельности	ОПК-3.1 Умеет анализировать задачу, планировать пути решения, предлагать и комбинировать способы решения

3. Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны знать:

математические, статистические методы и методы NLP, используемые на финансовых рынках.

уметь:

оценивать регрессионные модели, находить метрики качества, применять NLP к оценке тональности и построению финансовых стратегий.

владеть:

основами машинного обучения, используемыми в финансовых технологиях; техникой составления плана построения финансовых моделей.

4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

4.1. Разделы дисциплины (модуля) и трудоемкости по видам учебных занятий

№	Тема (раздел) дисциплины	Трудоемкость по видам учебных занятий, включая самостоятельную работу, час.			
		Лекции	Семинары	Лаборат. работы	Самост. работа
1	Анализ временных рядов.	5	5		6
2	Машинное обучение и нейронные сети в финансах.	5	5		6
3	Построение и тестирование моделей.	4	4		5
4	Оценка риска и бэк-тестинг моделей.	4	4		7
5	Построение стратегий оптимального инвестирования.	4	4		7
6	Основа NLP.	4	4		7
7	Применение NLP к финансовым задачам.	4	4		7
Итого часов		30	30		45
Подготовка к экзамену		30 час.			
Общая трудоёмкость		135 час., 3 зач.ед.			

4.2. Содержание дисциплины (модуля), структурированное по темам (разделам)

Семестр: 4 (Весенний)

1. Анализ временных рядов.

Временной ряд: примеры из жизни. Особенности предобработки (периодичность, сезонность, агрегация и дискретизация). Примеры, способы, задачи (выделение трендов, сезонных компонент, смена тренда, сумма временных рядов). Заполнение пропусков. Простейшие регрессии. Авторегрессии. AR, ARMA, ARIMA, GARCH, VAR, VECM и применение в финансах.

2. Машинное обучение и нейронные сети в финансах.

Decision Tree Regressor, Random Forest Regressor, XGBoost / LightGBM (особенности и примеры применения). FNN Recap, RNN Recap, GRU, LSTM, Transformers. Особенности для применения при работе с финансовыми временными рядами.

3. Построение и тестирование моделей.

Полный пайплайн: предобработка, агрегация, замена пропусков, выделение трендов и сезонностей, первый base-line и выбор архитектуры, метрики, сравнение результатов (несимметричная стоимость ошибки в + и -), TSCV (time series cross-validation), бэктестинг.

4. Оценка риска и бэктестинг моделей.

Оценка риска, Value at Risk, Expected Shortfall. Классические подходы к оценке риска актива на практике. Основные проблемы при оценке риска. Бэктестинг для определения качества оценки риска.

5. Построение стратегий оптимального инвестирования.

Методы оценки риска портфеля активов. Подходы с понижением размерности. Портфельная теория Марковица. Методы оптимизации портфеля.

6. Основа NLP.

Основы NLP (задачи, методы, инструменты). Токенизация, стемминг, лемматизация. Векторизация текста: TF-IDF, Word2Vec, FastText. Финансовый контекст. Особенности финансовых текстов: новости, отчеты, соцсети, регуляторные документы. Примеры применения NLP в финансах: прогнозирование рынка, анализ рисков, клиентский сервис. Инструменты и данные: библиотеки, источники данных, парсинг, агрегации, понятие качества и адекватности данных. Предобработка финансовых текстов. Распознавание именованных сущностей (NER). Обработка мультязычных данных.

7. Применение NLP к финансовым задачам.

Классификация финансовых текстов. Анализ тональности (Sentiment Analysis). Прогнозирование рынка на основе новостей. Извлечение ключевых данных. Вопросно-ответные системы (QA). Трансформеры в NLP (архитектура BERT, RoBERTa, FinBERT (модели, адаптированные для финансов, задачи: Masked Language Modeling, Next Sentence Prediction). Fine-tuning моделей под финансовые задачи.

5. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)

Учебная аудитория, оснащенная компьютером и мультимедийным оборудованием (проектор, звуковая система).

6. Перечень рекомендуемой литературы

Основная литература

1. Подготовка и защита бакалаврской работы, магистерской диссертации, дипломного проекта [Электронный ресурс], учеб. пособие / Ю. Н. Новиков. — СПб., Лань, 2019.— URL: <https://e.lanbook.com/book/122187> (дата обращения: 29.01.2021). - Полный текст (Режим доступа : из сети МФТИ / Удаленный доступ)

Дополнительная литература

7. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)

1. Questel Orbit <https://www.orbit.com/> – объединяет около 100 баз данных, предназначенных специалистам в области патентоведения и широкому кругу исследователей. Основная патентная база FamPat содержит данные 95 патентных ведомств всех регионов мира; патенты объединены в семьи по тематическому признаку.

2. Inspec Analytics – аналитический модуль базы данных Inspec <https://inspec-analytics-app.theiet.org/>. Inspec Analytics позволяет визуализировать результаты поиска, сравнивать полученные результаты на уровне учреждений, авторов, тематик по количеству публикаций.
3. Sage journals – более 100 журналов доступно в полнотекстовом режиме в области естественных наук, техники и медицины. <https://journals.sagepub.com/action/doSearch?filterOption=allJournal&AllField=research&content=journalTitle&target=titleSearch&pageSize=100&startPage=0>
4. Taylor&Francis journals – более 2000 журналов по всем областям знаний. Журналы разделены по коллекциям в области STM наук (Science, Technology & Medicine) и HSS (Humanities & Social Sciences), а также по более узким, конкретным областям знаний, <https://www.tandfonline.com/action/doSearch?AllField=research&startPage=&target=titleSearch&content=title>
5. Подготовка и защита бакалаврской работы, магистерской диссертации, дипломного проекта [Электронный ресурс], учеб. пособие / Ю. Н. Новиков. — СПб., Лань, 2019.— URL: <https://e.lanbook.com/book/122187> (дата обращения: 29.01.2021). - Полный текст (Режим доступа : из сети МФТИ / Удаленный доступ)
6. Искусство писать научные статьи, научно-практическое руководство / Е. З. Мейлихов. — Долгопрудный, Интеллект, 2020.— URL: <http://books.mipt.ru/book/301312> (дата обращения: 18.12.2020). - Полный текст (Режим доступа : из сети МФТИ / Удаленный доступ)

8. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень необходимого программного обеспечения и информационных справочных систем (при необходимости)

На лекционных занятиях используются мультимедийные технологии, включая демонстрацию презентаций.

В процессе самостоятельной работы обучающихся возможно использование таких программных средств, как Mathcad, MATLAB, Maple и др.

9. Методические указания для обучающихся по освоению дисциплины (модуля)

1. Рекомендуется успешно сдавать контрольные работы, так как это упрощает итоговую аттестацию по предмету.
2. Для подготовки к итоговой аттестации по предмету лучше всего пользоваться материалами лекций.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

по направлению: Прикладная математика и информатика
профиль подготовки: Комбинаторика и цифровая экономика
центр дополнительного, дополнительного профессионального и
онлайн-образования "Пуск"
кафедра дискретной математики
курс: 2
квалификация: магистр
Семестр, формы промежуточной аттестации: 4 (весенний) - Экзамен
Разработчик: А.В. Куликов, д-р физ.-мат. наук

1. Компетенции, формируемые в процессе изучения дисциплины

Код и наименование компетенции	Индикаторы достижения компетенции
УК-3 Способен организовывать и руководить работой команды, вырабатывая командную стратегию для достижения поставленной цели	УК-3.1 Организует и координирует работу участников проекта, способствует конструктивному преодолению возникающих разногласий и конфликтов
ОПК-2 Способен совершенствовать и реализовывать новые математические методы решения прикладных задач	ОПК-2.1 Имеет представление о современном состоянии математических исследований в рамках тематической области своей профессиональной деятельности
ОПК-3 Способен разрабатывать математические модели и проводить их анализ при решении задач в области профессиональной деятельности	ОПК-3.1 Умеет анализировать задачу, планировать пути решения, предлагать и комбинировать способы решения

2. Показатели оценивания компетенций

В результате изучения дисциплины «Финансовые технологии» обучающийся должен:

знать:

математические, статистические методы и методы NLP, используемые на финансовых рынках.

уметь:

оценивать регрессионные модели, находить метрики качества, применять NLP к оценке тональности и построению финансовых стратегий.

владеть:

основами машинного обучения, используемыми в финансовых технологиях; техникой составления плана построения финансовых моделей.

3. Перечень типовых (примерных) вопросов, заданий, тем для подготовки к текущему контролю

1. Портфельная теория Марковица.
2. Методы оптимизации портфеля.
3. Что такое NLP: задачи, методы, инструменты.
4. Токенизация, стемминг, лемматизация.
5. Векторизация текста: TF-IDF, Word2Vec, FastText.
6. Особенности финансовых текстов: новости, отчеты, соцсети, регуляторные документы.
7. Примеры применения NLP в финансах: прогнозирование рынка, анализ рисков, клиентский сервис.
- 8.. Библиотеки: spaCy, NLTK, Hugging Face Transformers.
- 9.. Источники данных. Парсинг. Агрегации. Понятие качества и адекватности данных.
- 10.. Очистка и нормализация
11. Удаление шума: HTML-теги, спецсимволы, стоп-слова.
12. Работа с финансовой терминологией: акронимы, тикеры акций (SBER).
13. Поиск компаний, индексов, валют, персон.
14. Использование предобученных моделей (spaCy, BERT) и дообучение на финансовых данных.
15. Особенности анализа текстов на разных языках, адаптация данных под доменную область с использованием трансформерных архитектур.
16. Категоризация новостей: слияния/поглощения, дивиденды, скандалы.
17. Основные алгоритмы: SVM, Random Forest, нейросети (LSTM). Техники ансамблирования. Выбор оптимального алгоритма.
18. Определение позитивного/негативного контекста в отчетах и соцсетях.
19. Финансовые лексиконы: Loughran-McDonald, SenticNet. Основные подходы к анализу тональности.

20. Fine-tuning моделей BERT для анализа настроений.
21. Прогнозирование рынка на основе новостей
22. Извлечение ключевых данных
23. Работа с таблицами и полуструктурированными данными (PDF, Excel).
24. Вопросно-ответные системы (QA)
25. Автоматизация ответов на запросы клиентов.
26. Использование моделей T5 и GPT-3.
27. Генерация текста
28. Автосоставление отчетов, пресс-релизов, инвестиционных рекомендаций.
29. Риски: этика и регуляторные ограничения.
30. Трансформеры в NLP
31. Архитектура BERT, RoBERTa, FinBERT (модели, адаптированные для финансов).

4. Перечень типовых (примерных) вопросов и тем для проведения промежуточной аттестации обучающихся

1. Временной ряд: примеры из жизни.
2. Периодичность и сезонность временных рядов.
3. Агрегация и дискретизация. Выделение трендов, сезонных компонент, смена тренда, сумма временных рядов
4. Заполнение пропусков (примеры и методы).
5. Регрессии и методы их калибровки.
6. AR, ARIMA, SARIMA, ARCH, GARCH.
7. VAR и VECM.
8. Decision Tree Regressor. Постановка и особенности.
9. Random Forest Regressor. Постановка и особенности.
10. XGBoost/LightGBM. Постановка и особенности.
11. FNN Recap. Постановка и особенности.
12. RNN Recap. Постановка и особенности.
13. Трансформеры. Постановка и особенности.
14. Предобработка, агрегация, замена пропусков, выделение трендов и сезонностей
15. Метрики качества модели в зависимости от постановки.
16. TSCV (time series cross-validation).
17. TSCV (time series cross-validation)
18. VaR, ESF.
19. Способы оценки VaR, ESF.
20. Бэктестинг VaR, ESF.
21. Методы оценки риска портфеля активов.
22. Подходы с понижением размерности

Билет 1:

1. Задачи: Masked Language Modeling, Next Sentence Prediction.
2. Fine-tuning моделей под финансовые задачи

Билет 2:

1. Дообучение BERT на SEC filings для прогнозирования волатильности.
2. Совместный анализ текста и числовых данных (например, квартальные отчеты + котировки).

Критерии оценивания

Оценка «отлично (10)» выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины и умение уверенно применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений;

Оценка «отлично (9)» выставляется студенту, показавшему систематизированные, глубокие знания учебной программы дисциплины и умение применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений;

Оценка «отлично (8)» выставляется студенту, показавшему систематизированные, знания учебной программы дисциплины и умение применять их на практике при решении конкретных задач, правильное обоснование принятых решений;

Оценка «хорошо (7)» выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач некоторые неточности;

Оценка «хорошо (6)» выставляется студенту, если он твердо знает материал, грамотно излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач некоторые неточности;

Оценка «хорошо (5)» выставляется студенту, если он знает материал, грамотно излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач некоторые неточности;

Оценка «удовлетворительно (4)» выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, недостаточно правильные формулировки базовых понятий, нарушения логической последовательности в изложении программного материала, но при этом он владеет основными разделами учебной программы, необходимыми для дальнейшего обучения и может применять полученные знания по образцу в стандартной ситуации;

Оценка «удовлетворительно (3)» выставляется студенту, показавшему фрагментарный характер знаний, недостаточно правильные формулировки базовых понятий, но при этом он владеет основными разделами учебной программы, необходимыми для дальнейшего обучения и может применять полученные знания по образцу в стандартной ситуации;

Оценка «неудовлетворительно (2)» выставляется студенту, который не знает большей части основного содержания учебной программы дисциплины, допускает грубые ошибки в формулировках основных понятий дисциплины и не умеет использовать полученные знания при решении типовых практических задач.

Оценка «неудовлетворительно (1)» выставляется студенту, который не знает основного содержания учебной программы дисциплины, допускает грубые ошибки в формулировках основных понятий дисциплины и не умеет использовать полученные знания при решении типовых практических задач.

5. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности

При проведении устного экзамена обучающемуся предоставляется 45 минут на подготовку.

Опрос обучающегося по билету не должен превышать одного астрономического часа.

Во время проведения экзамена обучающиеся могут пользоваться программой дисциплины.

ОТЧЕТ	
о научно-исследовательской работе за _____ семестр / _____ учебного года	
ФИО обучающегося	
Физтех-школа/факультет, группа	
Базовая организация, кафедра	
Тема НИР	
Текущее состояние НИР за семестр (проделанная работа и полученные результаты)	
И т о г и Н И Р з а с е м е с т Р	Доклады на научных конференциях, семинарах (авторы, название доклада и конференции, место проведения)
	Научные публикации (авторы, название работы и издания)
	Участие в конкурсах на лучшую НИР и выставках (авторы, название работы и конкурса (экспоната и выставки))
	Медали, дипломы, грамоты, премии и т.п. на конкурсах на лучшую НИР и на выставках (авторы, название работы и конкурса (экспоната и выставки), вид награды)
	Проекты, поданные на конкурсы грантов (авторы, название и вид гранта)
	Полученные гранты (авторы, название и вид гранта)
	Другое (заявки и охранные документы на объекты интеллектуальной собственности, проданные лицензии на их использование, стипендии Президента и Правительства РФ и т.п.)
Материальная поддержка НИР обучающегося за семестр (с указанием источника финансирования)	
План работы на следующий семестр	
Отзыв научного руководителя	
Рекомендованная оценка НИР обучающегося за семестр	

Обучающийся _____ дата составления отчета _____
 Научный руководитель _____ /ФИО/
 Зав. кафедрой _____ /ФИО/