КОНЦЕПЦИЯ МАГИСТЕРСКОЙ ПРОГРАММЫ
Финансовые технологии и анализ данных
НИУ ВШЭ

# Общая характеристика

Магистерская программа «Финансовые технологии и анализ данных» реализуется на Факультете компьютерных наук Высшей школы экономики (ФКН ВШЭ) совместно с ПАО «Сбербанк России». Она является практикоориентированной и обеспечивает как подготовку по информационным технологиям и современным методам анализа данных, так и ознакомление с основами финансового и банковского анализа, а также с теми прикладными задачами, которые возникают в современных финансовых data-driven организациях.

Форма обучения – очная.

Направление подготовки – 01.04.02 «Прикладная математика и информатика»

Набор студентов планируется осуществлять на платной основе. Тридцати поступающим, показавшим лучшие результаты по итогам вступительных испытаний и собеседования, Сбербанк оплачивает обучение на программе.

Нормативный срок освоения – 2 года, общая трудоемкость составляет 120 зачетных единиц.

Языки преподавания – русский и английский.

Выпускники программы получают документ об образовании и о квалификации государственного образца.

Партнером программы является ПАО «Сбербанк России», крупнейший банк России. ПАО «Сбербанк России» позиционирует себя как ИТ-компания с банковской лицензией и является крупнейшим инвестором в информационные технологии в России[[1]](#footnote-1).

# Актуальность и значимость

Международный лидер в сфере услуг стратегического консалтинга, компания McKinsey, публикует исследования, согласно которым к 2025 году мировая банковская индустрия претерпит существенные изменения[[2]](#footnote-2). Одними из главных причин трансформации являются диджитализация банковских услуг, стремительно растущий объем данных о клиентах и их операциях, возникновение новых видов риска, связанных с использованием машинного обучения в принятии решений.

Данные тенденции определяют высокий спрос на редкую комбинацию знаний и профессиональных навыков у специалистов, желающих трудоустроиться в банковской сфере. Дата-аналитик должен обладать знаниями об устройстве финансовой системы, макроэкономике, бизнес-процессах и при этом владеть навыками сбора, обработки и анализа больших массивов данных, а также уметь развивать и применять современные методы математического моделирования и предиктивной аналитики.

В ближайшем будущем информационные технологии настолько сильно изменят ландшафт финансовой системы, что владение данными технологиями и их успешное внедрение в бизнес-процессы будет не просто определять лидерство на рынке, но само выживание на этом рынке. Данное опасение было высказано генеральным директором такого крупного международного банка как JP Morgan[[3]](#footnote-3). В выступлениях руководителей различных банков фигурируют такие высказывания, как «Большие данные – это новая нефть», «Машинное обучение – это новое электричество», которые подчеркивают, что анализ больших массивов данных является не просто конкурентным преимуществом, а жизненно необходимым элементом любой крупной финансовой организации.

Развитие информационных технологий создает гораздо более жесткую конкурентную среду для банков и кредитных организаций. Так, например, с января 2017 г. две крупнейшие телекоммуникационные компании России начали предоставлять займы для своих клиентов[[4]](#footnote-4), хотя ранее никогда не развивали данное бизнес-направление. Предоставление подобных услуг со стороны нефинансовых компаний оказалось возможным именно благодаря внедрению современной ИТ-инфраструктуры для хранения большого объема данных по клиентам и применению продвинутых математических методов анализа данных, таких как машинное обучение. Выходя на рынок кредитования, нефинансовые компании, естественно, претендуют на долю рынка, конкурируя с традиционными участниками банковского сектора.

Другим примером технологии, которая может радикально изменить функционирование банковской системы, является активно развивающаяся технология блокчейн (block-chain). В середине 2016 г. глава ПАО «Сбербанк России» Г.О.Греф высказал следующие ожидания: «Я считаю блокчейн исключительно перспективной технологией. Через два-три года мало останется сфер, где бы она радикально не изменила основы бизнес-модели»[[5]](#footnote-5).

Приведенные выше примеры показывают, что изменения, связанные с информационными технологиями и современными математическими методами анализа данных, окажут существенное влияние на финансовую отрасль, и, при недостаточном уровне подготовки участников финансовой системы, могут стать угрозой для их успешного функционирования.

Одним из главных участников финансовой системы российской экономики является ПАО «Сбербанк России» (Сбербанк). Сбербанк сегодня — это треть российской банковской системы. Таким образом, современные тренды в развитии информационных технологий и математических методов анализа данных становятся вызовом не только перед Сбербанком как отдельной крупной компанией, но и всей финансовой системой в России.

Ответом на данный вызов, прежде всего, должна стать подготовка высококвалифицированных специалистов, которые обладают знаниями в области макроэкономики, финансов, управления бизнес-процессами и одновременно владеют навыками анализа больших массивов данных, разработки математических моделей и предиктивной аналитики. На данный момент в нашей стране этот вызов ещё не встретил должного ответа со стороны образовательного сообщества.

# Цели магистерской программы

Целью создания магистерской программы «Финансовые технологии и анализ данных» является ответ основным вызовам, обращённым как к НИУ ВШЭ и российскому образованию в целом, так и к российским системообразующим финансовым организациям: подготовка высококвалифицированных специалистов, обладающих уникальной комбинацией компетенций в передовых областях развития информационных технологий, математики и экономики.

 Обучающиеся проходят фундаментальную подготовку по релевантным разделам математики, требующимся для анализа данных, осваивают современные парадигмы программирования (параллельно-распределенные вычисления, cloud computing), современные методы машинного обучения к задачам финансов, а также изучают основы макроэкономики, корпоративных финансов и банковского дела.

Особое внимание в рамках магистерской программы уделяется применению полученных знаний на практике. Курсы включают разбор прикладных задач, связанных с анализом данных, которые возникают в крупных финансовых data-driven компаниях. В этой перспективе ключевую роль играет партнер магистерской программы – Сбербанк. Планируется предоставление студентам возможности изучать курсы для руководителей банка, реализуемые в Корпоративном университете Сбербанка; привлечение сотрудников банка к преподаванию и проведение мастер-классов с их участием; разбор кейсов и проведение «хакатонов»; работа над проектами на основе массивов данных, предоставляемых банком; проведение производственной практики в подразделениях банка с организацией обзорных выступлений руководителей этих подразделений о их задачах и текущей работе.

Экономические и финансовые дисциплины будут представлены как в адаптационных, так и в базовых обязательных курсах программы.

Практическая направленность программы будет способствовать повышению конкурентоспособности и востребованности выпускников программы.

Основными целями магистерской программы являются:

1. Получение студентами, имеющими базовое образование в области прикладной математики, информатики, экономики и смежных областях, специализации магистерского уровня в сфере анализа данных, машинного обучения, искусственного интеллекта с приложением к задачам в области финансов.
2. Дополнение аналитических компетенций, полученных при изучении экономики, эконометрики, математической статистики, компьютерной лингвистики и т.п., математическими и технологическими компетенциями по работе с большими данными, различной структуры.
3. Получение выпускниками, имеющими образование в сфере анализа данных, специализации в области финансов банка и основных областей применения анализа данных – финансовых технологий, риск-менеджмента, обработки данных клиентов, мобильных приложений и др.

## Целевая аудитория представляет собой следующие категории:

1. Выпускники бакалавриата математических, экономических и технических факультетов ведущих университетов.
2. Специалисты и магистры по направлениям, указанным в п.1, имеющие целью получить дополнительную специализацию и системно подготовиться к серьёзной аналитической или исследовательской деятельности в области анализа больших данных в современных крупных финансовых компаниях.

# Требования к абитуриентам и условия поступления

Поступающие на магистерскую программу «Финансовые технологии и анализ данных» должны иметь высшее образование (степень бакалавра и/или специалиста), подтвержденное документом государственного образца.

На момент поступления необходимо владеть английским языком на уровне достаточном для освоения научной, технической и учебной литературы, а также лекций, читаемых на английском языке.

Также на момент поступления необходимо обладать достаточными знаниями в области:

* Математического анализа,
* Линейной алгебры,
* Дискретной математики,
* Теории вероятности,
* Математической статистики,
* Многомерного статистического анализа,
* Алгоритмов и структур данных.

Также поступающие должны обладать навыками программирования. Прием на магистерскую программу осуществляется путем конкурсного отбора на основе результатов вступительных испытаний и собеседования. Таким образом, конкурсный отбор состоит из трех этапов:

1. Экзамен по направлению подготовки 01.04.02 «Прикладная математика и информатика», включающий задачу на основы программирования;
2. Экзамен по английскому языку;
3. Собеседование.

Для подтверждения необходимого уровня владения английским языком, кандидаты могут представить международный сертификат экзамена по английскому языку САЕ (Certificate of Advanced English); ВЕС Higher (Business English Certificate); IELTS с оценкой не ниже 6, TOEFL IBT (Internet Based) от 80 баллов (рассчитывается от 0 до 120) или PBT (Paper Based) – от 500 баллов (рассчитывается от 310 до 677). В случае отсутствия международного сертификата кандидат сдает квалификационный экзамен по английскому языку.

По результатам вступительных испытаний и собеседования поступающим, показавшим лучшие результаты, Сбербанк оплачивает обучение на программе без обязательств со стороны поступающего в отношении работы в Сбербанке после окончания программы. Количество мест, финансируемых Сбербанком, составляет 30 (тридцать) мест. В случае набора свыше 30 обучающихся студенты, не вошедшие в 30 лучших по результатам вступительных испытаний и собеседования, оплачивают обучение самостоятельно.

# Характеристика сегмента рынка образовательных услуг в сфере анализа данных и основные конкуренты

Учебные курсы, посвященные данной тематике, в России являются либо частью образовательных программ в университетах, либо читаются по инициативе компаний, либо развиваются открытыми онлайн сообществами.

Среди магистерских программ, которые близки данному направлению, отметим программы «Науки о данных», «Статистическая теория обучения» на факультете компьютерных наук НИУ ВШЭ, программу «Анализ данных» ФУПМ МФТИ, программы «Интеллектуальный̆ анализ больших данных» и «Статистический анализ и прогнозирование рисков» ВМК МГУ, «Математические и компьютерные методы анализа» мехмата МГУ, программу «Управление большими данными» в университете Иннополис, а также «Master Programmes in Applied Mathematics and Stochastics» в НГУ, программу «Финансовая математика» на мехмате МГУ.

Среди образовательных инициатив компаний, которые близки данному направлению, отметим Школу анализа данных компании Яндекс (<https://yandexdataschool.ru/>), Школу данных компании Билайн (<http://bigdata.beeline.digital>), открытие кафедры финансовых технологий компании Тинькофф в МФТИ.

Среди открытых онлайн сообществ отметим сообщество Open Data Science (<http://ods.ai/>), курсы на платформе Coursera (<https://www.coursera.org/course/datasci>).

Из международного опыта магистерских программ отметим Университет Нью-Йорка (<http://cds.nyu.edu/academics/ms-in-data-science/curriculum>), Вашингтона (<http://www.pce.uw.edu/certificates/data-science.html>), Карнеги-Меллон (<http://www.heinz.cmu.edu/school-of-information-systems-and-management/information-systems-management-mism/business-intelligence-data-analytics/index.aspx>), программу в Калифорнийском университете в Беркли (<http://datascience.berkeley.edu>) а также программу MSc Business Analytics в Imperial College Business School[[6]](#footnote-6).

Большинство университетских программ имеют либо академический акцент на преподавание дисциплин (выпускники подготавливаются скорее для исследовательской и научной работы), либо не фокусируются на приложениях в финансовой сфере.

Образовательные программы от компаний сфокусированы на прикладные аспекты изучения анализа данных и не предоставляют степень магистра. Кроме того, отметим, что сферой деятельности компаний Яндекс и Билайн являются ИТ и телекоммуникации соответственно – предоставление финансовых продуктов и услуг не является их приоритетной деятельностью, в отличие от Сбербанка. Финтех школа банка Тинькофф представляет собой обучающий курс по языку программирования Scala и его приложениям в задачах анализа финансовых данных длительностью всего два месяца. Образовательная программа компании WorldQuant так же предполагает финансовую направленность, но является узконаправленной – рассматриваются методы построения математических моделей исключительно для трейдинга на фондовом рынке.

Образование в рамках открытых онлайн сообществ имеет целью, прежде всего, обмен опытом и не предполагает каких-либо вступительных испытаний, получения степени и служит, скорее, средством коммуникации между профессионалами и начинающими специалистами в области анализа данных и машинного обучения.

# Анализ востребованности выпускников и возможности трудоустройства

Подтверждением потребности в магистрах данного профиля служат многочисленные консалтинговые компании, оказывающие услуги в области анализа данных, среди них отметим такие представленные в России компании, как SAS, IBM, Accenture, Oliver Wyman, KPMG, Deloitte, Glowbyte, Double Data, Rubbles, Econophysica, Prognoz, Forecsys, и др. Все они предъявляют спрос на высококвалифицированных специалистов в области математической статистики, машинного обучения и анализа данных. Сбербанк за последние 7 лет имел опыт работы практически со всеми данными консультантами, и является своего рода аккумулятором экспертизы.

В свою очередь сам Сбербанк активно нанимает специалистов в данной сфере. На данный момент в Сбербанке трудится более 120 дата-аналитиков, планируется расширение их численности, а также предполагается трансформация банка в data-driven организацию. Выпускники программы будут востребованы на позициях специалистов и руководителей проектов в подразделениях, отвечающих за сбор, обработку, анализ больших массивов разнородных данных, построение математических моделей на основе методов машинного обучения, имплементацию математических моделей в промышленных средах, мониторинг качества работы моделей, разработку продуктов и сервисов, основанных на анализе больших массивов данных.

Банк проводит «хакатоны» по анализу данных, на которые предоставляет реальные данные из своих систем – например, Sberbank Data Science Journey[[7]](#footnote-7).

Классическая схема подготовки аналитиков не соответствует этим вызовам, поскольку системно не охватывает дополнительные задачи обработки и анализа данных в финансовой сфере, включая неструктурированные данные больших объёмов, а также не уделяет должного внимания практическому опыту использования современных методов хранения и обработки данных (Apache Spark, Hadoop, Hive, MongoDB). При этом присутствует дефицит специалистов, готовых системно подходить к решению задач, связанных с обработкой данных разных видов и типов, перестройкой структуры хранилищ данных, эффективностью процессов обработки, анализом больших данных (снижение размерности, обнаружение аномалий, схема проведения статистических экспериментов и т.д.) применительно к финансовой сфере.

# Набор компетенций, формируемых программой

Магистерская программа формирует у обучающихся следующие навыки и компетенции:

1. Применяет на практике знания о макроэкономических аспектах устройства финансовой системы, финансах банков, банковской ИТ-инфраструктуре, об основных задачах анализа данных в области финансовых технологий, управления рисками, банковском маркетинге (обработка данных клиентов).
2. Строит математические модели и вычислительные алгоритмы обработки и анализа данных, принятия решений в банковской сфере с учетом неопределенности и риска, оценивает качество работы таких алгоритмов и интерпретирует результаты моделирования (с использованием таких инструментов, как Python, R, H2O, Spark MLLib).
3. Применяет на практике системные знания о жизненном цикле проектов и об инструментах управления проектами в сфере ИТ.
4. Формулирует бизнес-задачи в контексте сбора необходимых данных, анализа данных, построения математической модели, имплементации в промышленной среде.
5. Использует на практике современные технологии обработки больших объёмов сложно или недостаточно структурированных данных (такие как Apache Spark, Storm, Kafka, Cassandra).

# Особенности структуры и содержания программы

## Оригинальный образовательный стандарт

Магистерская программа опирается на оригинальный образовательный стандарт (ОрОС) НИУ ВШЭ по направлению 01.04.02 «Прикладная математика и информатика».

Стандарт включает теоретическое обучение – цикл дисциплин направления и цикл дисциплин программы. В стандарт входят:

1. Научно-исследовательский семинар.
2. Междисциплинарная курсовая работа.
3. Научно-исследовательская практика.
4. Итоговая государственная аттестация.
	1. Междисциплинарный экзамен по направлению «Прикладная математика и информатика».
	2. Подготовка и защита выпускной квалификационной работы.

Поскольку целевой аудиторией предполагаются выпускники разнородных бакалаврских программ различных университетов, в 1 модуле 1 курса предусмотрены адаптационные дисциплины, включающие курсы по экономике.

## Базовый учебный план

Адаптационные курсы.

1. Макроэкономика
2. Теория баз данных
3. Эконометрика
4. Машинное обучение

Цикл дисциплин направления.

Базовая часть:

1. Современные методы анализа данных: глубинное обучение.
2. Современные методы принятия решений: алгоритмы обработки больших данных

Вариативная часть:

1. Финансы банка
2. Финансовые технологии

Цикл дисциплин программы.

Базовая часть

1. Управление данными и исполнение моделей
2. Байесовские методы в анализе данных
3. Анализ текстов. Генеративные модели.
4. Основы риск-менеджмента
5. Банковская ИТ-инфраструктура

Вариативная часть

1. Архитектура кластерных систем и параллельно-распределенное выполнение
2. Рекомендательные системы и поиск закономерностей в данных
3. Технологии мобильных приложений
4. Управление ИТ-проектами и командами разработчиков в сфере data science
5. Структурный анализ и визуализация сетей
6. Онтологическое моделирование и семантические технологии

## Концепция научно-исследовательского семинара

Важным элементом магистерской программы является Научно-исследовательский семинар (НИС). Проведение НИС преследует две цели:

1. Подготовить магистранта к культуре воспроизводимого исследования (reproducible research), которая включает в себя умение изложить свои результаты, предоставить средства к воспроизведению результатов, понять и изложить чужие результаты, а также умение принять участие в дискуссии по рабочим вопросам;
2. Подготовить магистранта к защите курсовой работы или ВКР практической направленности, например, решение практического бизнес-кейса в финансовой сфере с использованием методов машинного обучения, разработка новых алгоритмов анализа данных и исследование их качества работы.

Достижение вышеуказанных целей позволит подготовить студента к участию в бизнес-совещаниях, переговорах и защите своей позиции. Программа семинара включает три основных вида деятельности:

1. подготовка и проведение доклада о собственной курсовой работе или ВКР;
2. изучение материалов и подготовка доклада об опубликованной в международной научной литературе работе;
3. активное участие в обсуждении работы коллег.

Оцениваются все три вида деятельности. Для повышения эффективности обучения предоставляется дополнительная информация по подготовке презентаций, правилам ведения дискуссии и т.п.

НИС может проводиться на различных площадках, включая корпоративные. Его главная цель – поддержать прикладную научно-исследовательскую деятельность студента.

# Ресурсное обеспечение

## Обеспеченность профессорско-преподавательским составом

ФКН ВШЭ является ведущим центром компетенций в области анализа данных и машинного обучения, нейросетевых технологий, оптимизации и стохастики, теоретической информатики, системного программирования и других разделов современных компьютерных наук. Научно-практическая деятельность на факультете развивается в таких лабораториях, как Лаборатория методов анализа больших данных (зав. Устюжанин А.Е.), международная Лаборатория глубинного обучения и байесовских методов (зав. Ветров Д.П.), международная Лаборатория интеллектуальных систем и структурного анализа (зав. С.О.Кузнецов). На факультете работают ведущие специалисты в области анализа данных, майнинга данных и интеллектуальных систем. Среди них отметим профессора д.ф.-м.н. Нестерова Ю.Е., ведущего научного сотрудника МНУЛ Анализа и выбора решений д.т.н. Миркина Б.Г., зав. НУЛ Процессно-ориентированных информационных систем д.ф.-м.н. Ломазову И.А., заместителя руководителя Департамента больших данных и информационного поиска Соколова Е.А.

Академическое управление программой будет осуществляться академическим руководителем и академическим советом. Руководитель магистерской программы ­– Масютин А.А., специалист в области математических моделей и кредитного скоринга, руководитель проектного офиса Управления валидации блока Риски в Сбербанке. В Академический совет программы, возглавляемый старшим вице-президентом ПАО «Сбербанк России» Ведяхиным А.А., войдут представители НИУ ВШЭ, Корпоративного университета и сотрудники различных подразделений ПАО «Сбербанк России».

Ядром профессорско-преподавательский состава (ППС) магистерской программы является ППС департамента анализа данных и искусственного интеллекта, департамента больших данных и информационного поиска, базовой кафедры Яндекс и других подразделений факультета компьютерных наук НИУ ВШЭ. К проведению занятий в рамках программы также планируется привлечь сотрудников Корпоративного университета Сбербанка, а также специалистов Сбербанка в качестве авторов мастер-классов, участников разбора бизнес-кейсов, руководителей курсовых и выпускных работ.

Среди сотрудников Корпоративного университета Сбербанка – разработчиков и преподавателей ряда дисциплин данной программы:

* Баранов И.Н., к.э.н., доцент, МВА (London Business School), МРА (Columbia University), проректор КУ;
* Волков Д.Л., д.э.н., профессор, директор по образовательным технологиям КУ;
* Удовиченко О.М., к.э.н., доцент, МВА (Duke University), проректор, руководитель Школы риск-менеджмента КУ;
* Щербакова О.Н., к.э.н., руководитель Школы финансов КУ (также - преподаватель финансов РЭШ).
* Шибанов О.К., Ph.D. in Economics (London Business School) (также - профессор финансов РЭШ);

Со стороны специалистов Сбербанка в преподавательской и научно-практической деятельности будут задействованы сотрудники блока Риски, Центров компетенций по моделированию, анализа супермассивов данных. Эти сотрудники имеют практический опыт успешного использования подходов математического моделирования, методов машинного обучения, разработки информационных систем и платформ в банковской сфере:

* Марков Е.А. – к.ф.-м.н. руководитель направления контроля и подготовки данных в управлении валидации моделей блока Риски;
* Рузанов Д.П. – начальник отдела валидации моделей розничного бизнеса, сотрудник Лаборатории анализа финансовых рынков НИУ ВШЭ;
* Загорулькин Д.Э – руководитель разработки в Центре компетенций по анализу супермассивов данных в Сбертехе;
* Крайнов Д.Е. – управляющий директор, автор более 10 публикаций по риск-менеджменту и вопросам банковского регулирования.

## Обеспеченность оборудованием

НИУ ВШЭ укомплектована необходимыми средствами для проведения занятий с использованием компьютеров: компьютерными классами с современными рабочими станциями. Также возможна поддержка со стороны Сбербанка с предоставлением доступа к лабораторному вычислительному кластеру.

1. <http://www.rbc.ru/newspaper/2015/06/24/56bcc4ea9a7947299f72beb0> [↑](#footnote-ref-1)
2. [http://www.mckinsey.com/~/media/mckinsey/business%20functions/risk/our%20insights/the%20future%20of %20bank%20risk%20management/the-future-of-bank-risk-management-full-report.ashx](http://www.mckinsey.com/~/media/mckinsey/business%20functions/risk/our%20insights/the%20future%20of%20%20bank%20risk%20management/the-future-of-bank-risk-management-full-report.ashx) [↑](#footnote-ref-2)
3. <http://mashable.com/2015/04/10/jp-morgan-ceo-letter/#cmdLF08POPqy> [↑](#footnote-ref-3)
4. <http://www.rbc.ru/finances/11/01/2017/587500529a794767fa723fa8> [↑](#footnote-ref-4)
5. <https://lenta.ru/news/2016/06/08/blockgref/> [↑](#footnote-ref-5)
6. [http://wwwf.imperial.ac.uk/business-school/programmes/msc-business-analytics/?gclid=Cj0KEQjwqtjG BRD8yfi9h42H9YUBEiQAmki5Ola384AviINQSQVHNOCnkKxOpmJwLaHK-HAP63s0HRAaAi718P8HAQ](http://wwwf.imperial.ac.uk/business-school/programmes/msc-business-analytics/?gclid=Cj0KEQjwqtjG%20BRD8yfi9h42H9YUBEiQAmki5Ola384AviINQSQVHNOCnkKxOpmJwLaHK-HAP63s0HRAaAi718P8HAQ) [↑](#footnote-ref-6)
7. <https://sdsj.ru/> [↑](#footnote-ref-7)