



Национальный исследовательский университет «Высшая школа экономики»
Программа дисциплины [Введите название дисциплины] для направления/ специальности [код направле-
ния подготовки и «Название направления подготовки»] подготовки бакалавра/ магистра/ специалиста

**Санкт-Петербургский филиал федерального государственного
автономного образовательного учреждения высшего образования
«Национальный исследовательский университет
"Высшая школа экономики"»**

Факультет Санкт-Петербургская школа
физико-математических и компьютерных наук

Департамент математики

**Рабочая программа дисциплины
«Глубокое обучение»**

для образовательной программы «Анализ больших данных в бизнесе, экономике и
обществе» направления подготовки 01.04.02 «Прикладная математика и информатика»
уровень магистратура

Разработчики программы
Сироткин А.В., avsirotkin@hse.ru
Николенко С.И., snikolenko@hse.ru

Согласована начальником ОСУП
«31» августа 2018г.

О.Н. Колобова _____

Утверждена Академическим руководителем образовательной программы
«31» августа 2019г.

А.В. Сироткин _____



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

1 Область применения и нормативные ссылки

Настоящая программа учебной дисциплины устанавливает требования к образовательным результатам и определяет содержание и виды учебных занятий и отчетности.

Программа предназначена для преподавателей, ведущих дисциплину «Глубокое обучение», учебных ассистентов и студентов направления подготовки 01.04.02 «Прикладная математика и информатика», обучающихся по образовательной программе «Анализ больших данных в бизнесе, экономике и обществе».

Рабочая программа дисциплины разработана в соответствии с:

- Образовательным стандартом НИУ ВШЭ, утвержденным ученым советом Национального исследовательского университета «Высшая школа экономики», протокол от 22.12.2017 №13

<https://www.hse.ru/data/2018/02/09/1162000644/01.04.02%20%D0%9F%D1%80%D0%B8%D0%BA%D0%BB%D0%B0%D0%B4%D0%BD%D0%B0%D1%8F%20%D0%BC%D0%B0%D1%82%D0%B5%D0%BC%D0%B0%D1%82%D0%B8%D0%BA%D0%B0%20%D0%B8%20%D0%B8%D0%BD%D1%84%D0%BE> образовательной программой «Анализ больших данных в бизнесе, экономике и обществе», направление подготовки 01.04.02 «Прикладная математика и информатика»;

- Объединенным учебным планом университета по образовательной программ «Анализ больших данных в бизнесе, экономике и обществе».

2 Цели освоения дисциплины

Целями освоения дисциплины «Глубокое обучение» являются

- формирование теоретических знаний о математических методах поиска и анализа данных для принятия и реализации решений.
- формирование представление о системах поддержки принятия решений.

3 Компетенции обучающегося, формируемые в результате освоения дисциплины

В результате освоения дисциплины студент осваивает следующие компетенции:

Компетенция	Код по ОС ВШЭ	Дескрипторы – основные признаки освоения (показатели достижения результата)	Формы и методы обучения, способствующие формированию и развитию компетенции	Форма контроля уровня сформированности компетенции
Способен рефлексировать (оценивать и перерабатывать) освоенные научные методы и способы деятельности.	УК-1	Способен выбирать подходящие структуры нейронных сетей.	Практические занятия, самостоятельная работа, лекции	Контрольная работа, экзамен



Способен анализировать, верифицировать, оценивать полноту информации в ходе профессиональной деятельности, при необходимости восполнять и синтезировать недостающую информацию.	УК-6	Способность находить и самостоятельно осваивать нужную информацию из документации, литературы, сети Интернет	Практические занятия, самостоятельная работа, лекции	Контрольная работа, экзамен
Способен применять системный подход при постановке задач и выборе подходов к решению, а также для учёта противоречивых целей, потребностей и требований.	ОПК-1	Способен оценивать скорость обучения нейронной сети и качество этого обучения. Понимать, когда обучение стоит прервать.	Практические занятия, семинарские занятия, самостоятельная работа, лекции	Контрольная работа, экзамен
Способен обоснованно выбирать и применять в профессиональной деятельности современные компьютерные технологии в соответствии со спецификой решаемых задач, включая операционные системы, сетевые технологии, языки программирования, языки манипулирования данными, электронные библиотеки, пакеты прикладных программ.	ОПК-3	Способен выбирать существующие реализации нейронных сетей и фреймворки в соответствии с требованиями задачи.	Практические занятия, семинарские занятия, самостоятельная работа, лекции	Контрольная работа, экзамен
Способен строить математические модели и использовать их при решении прикладных задач в соответствии с направлением подготовки и специализацией	ОПК-5	Способен эффективно выбирать режимы обучения нейронных сетей.	Практические занятия, семинарские занятия, самостоятельная работа, лекции	Домашняя работа, экзамен



Способен получать, очищать, анализировать и визуализировать большие объёмы данных	ПК-9	Способен строить сложные модели обработки текста и изображений с помощью глубоких нейронных сетей.	Практические занятия, семинарские занятия, самостоятельная работа, лекции	Контрольная работа, домашняя работа, экзамен
Способен реализовывать модели и алгоритмы прикладной математики в виде компьютерных программ.	ПК-10	Может создать программную реализацию заданной структуры нейронной сети.	Практические занятия, семинарские занятия, самостоятельная работа, лекции	Контрольная работа, экзамен

4 Место дисциплины в структуре образовательной программы

Дисциплина «Современные методы принятия решений» относится к базовой части цикла дисциплин направления 01.04.02 «Прикладная математика и информатика».

Изучение данной дисциплины базируется на следующих дисциплинах:

- «Современные методы анализа данных»,
- «Машинное обучение»
- «Алгоритмы и структуры данных»,
- «Практическое программирование и анализ данных в специализированных средах».

Для освоения учебной дисциплины, студенты должны владеть следующими знаниями и компетенциями:

- Способен создавать новые теории, изобретать новые способы и инструменты профессиональной деятельности
- Способен совершенствовать и развивать свой интеллектуальный и культурный уровень, строить траекторию профессионального развития и карьеры.
- Способен публично представлять результаты профессиональной деятельности (в том числе с использованием информационных технологий).
- Способен описывать проблемы и ситуации профессиональной деятельности, используя язык и аппарат прикладной математики при решении междисциплинарных проблем
- Способен применять в исследовательской и прикладной деятельности современные языки программирования и языки манипулирования данными, операционные системы, электронные библиотеки и пакеты программ, сетевые технологии и т.п.

Основные положения дисциплины могут быть использованы в дальнейшем при подготовке ВКР и обучении в аспирантуре.

5 Тематический план учебной дисциплины

Курс рассчитан на 64 часа аудиторной нагрузки, из них 32 часа лекций и 32 часа семинаров, общим объемом 6 зачетных единиц (228 часов).

№	Название раздела	Всего часов	Аудиторные часы		Самостоятельная работа
			Лекции	Семинары	
1	Графические модели	58	10	12	36



2	Основы нейронных сетей	50	8	6	36
3	Специальные виды нейронных сетей	52	8	8	36
4	Выбор стратегий с помощью нейросетей	68	6	6	56
ИТОГО		228	32	32	164

6 Содержание дисциплины

1. Раздел 1 Графические модели

Графические модели: определения, обозначения, примеры.

Маргинализация в общем виде, вывод на графе без циклов.

Вывод на графе с циклами: вариационные приближения. Алгоритм EM в общем виде.

Сэмплирование как метод приближённого вычисления. Методы сэмплирования.

Тематическое моделирование и модель LDA.

Вывод в моделях со сложными факторами: Expectation Propagation.

Байесовские рейтинг-системы.

2. Раздел 2 Основы нейронных сетей

Нейронные сети: перцептрон. Виды функций активации.

Обучение одного перцептрона. История развития нейронных сетей.

Градиентный спуск. Обратное распространение градиента на графе вычислений.

Как сделать градиентный спуск быстрее и лучше.

Метод моментов, методы второго порядка и другие трюки.

Регуляризация в нейронных сетях. Дропаут и его мотивация. Другие методы.

3. Раздел 3 Специальные виды нейронных сетей

Рекуррентные сети: базовые архитектуры, LSTM, GRU.

Свёрточные сети: архитектуры, как обучать, для чего они нужны.

Глубокие сети для обработки текстов I: распределённые представления слов.

Глубокие сети для обработки текстов II: рекурсивные нейронные сети, сети со стеком, сети с памятью.

4. Раздел 4 Выбор стратегий с помощью нейросетей

Обучение с подкреплением: основы, определения, классические алгоритмы

Как работает AlphaGo: обучение с подкреплением на глубоких сетях. DQN.

Соединяем байесовский вывод и глубокие сети: нейробайесовские методы.

7 Оценочные средства

7.1. Формы контроля знаний студентов

Тип контроля	Форма контроля	2 курс		Параметры
		1 модуль	2 модуль	
Текущий	Контрольная работа	*		Письменная работа 80 минут
	Домашнее задание		*	
Итоговый	Экзамен		*	Письменный экзамен 150 мин.



7.2. Критерии и шкалы оценки, примеры заданий

7.2.1. Текущий контроль

Текущий контроль успеваемости по дисциплине осуществляется с помощью следующих оценочных средств.

Контрольная работа

Контрольная работа представляет собой набор заданий (с подзаданиями). В конце контрольной работы результаты сдаются в виде кода с комментариями и ответами на поставленные вопросы.

Пример задания на контрольную работу:

1. Для заданной структуры байесовской сети доверия, укажите условно независимые переменные, если в узел X поступило свидетельство.
2. Чем отличается стохастический градиентный спуск от обычного.
3. Опишите метод обратного распространения ошибки на примере сети с 3-мя слоями, состоящими из двух узлов каждый.

Критерии оценивания и шкала оценки контрольной работы

Оценка	Критерии выставления оценки
«Отлично» (8-10)	Ответы на все вопросы даны полном объеме. Нет явных ошибок.
«Хорошо» (6-7)	Даны ответы на все или все кроме одного задания. Допускается неточности в отдельных местах, но нет существенных ошибок.
«Удовлетворительно» (4-5)	Ответы на вопросы даны частично, но в объеме позволяющем проверить корректность ответов.
«Неудовлетворительно» (0-3)	Ни на один вопрос не дан содержательный ответ.

Домашнее задание

Домашнее задание выдается студентам в одном варианте. Срок выполнения домашнего задания 2 недели. Решение необходимо представить в электронном виде, или в виде кода с комментариями или в виде отдельного отчета о результатах, с приложением кода. Программная реализация допустима на языке, согласованном с преподавателем, по умолчанию весь код пишется на Python.

Пример домашнего задания

Для набора данных MNIST обучите нейронную сеть. Сравните качество вашей сети с предложенными ранее методами.

Критерии оценивания и шкала оценки домашней работы

Оценка	Критерии выставления оценки
«Отлично» (8-10)	Задание домашней работы выполнено в полном объеме. Нет явных ошибок. Код воспроизводит те же результаты, которые даны в отчете.



«Хорошо» (6-7)	Выполнена существенная часть задания. Отчет не содержит явных ошибок и проведен качественный анализ полученных результатов.
«Удовлетворительно» (4-5)	Задание выполнено частично, но в объеме позволяющем проверить корректность кода и полученных результатов.
«Неудовлетворительно» (0-3)	Задание не выполнено или выполнено с грубыми ошибками. Код не запускается или при применении кода к исходным данным результаты отличаются от описанных в работе.

7.2.2. Итоговый контроль

Итоговый контроль проходит в форме письменного устного экзамена продолжительностью 150 минут.

Пример задания на письменный экзамен.

1. Нарисуйте схему LSTM сети.
2. Объясните принципы обучения с подкреплением на примере модели многорукого бандита.
3. Что такое условная независимость и какова ее роль в вероятностных графических моделях.

Критерии оценивания и шкала оценки экзамена

Оценка	Критерии выставления оценки
«Отлично» (8-10)	Ответы на все вопросы даны полном объеме. Нет явных ошибок.
«Хорошо» (6-7)	Даны ответы на все или все кроме одного задания. Допускается неточности в отдельных местах, но нет существенных ошибок.
«Удовлетворительно» (4-5)	Ответы на вопросы даны частично, но в объеме позволяющем проверить корректность ответов.
«Неудовлетворительно» (0-3)	Ответ отсутствует полностью или демонстрирует непонимание студентом предмета.

7.3 Порядок формирования оценок по дисциплине

Результующая оценка по дисциплине (которая идет в диплом) рассчитывается следующим образом:

$$O_{\text{накопленная}} = 0,36O_{\text{дз}} + 0,24O_{\text{кр}} + 0,4O_{\text{экз}}, \text{ где}$$

$O_{\text{дз}}$ - оценка знаний студента за домашнее задание;

$O_{\text{кр}}$ – оценка знаний студента за контрольную работу;

$O_{\text{экз}}$ - оценка за экзамен.



В формулу для **О_{результ}** подставляются значения **О_{накопленная}** и **О_{экза}**, округленные до ближайшего целого значения (по арифметическим правилам). **О_{результ}** округляется до ближайшего целого значения (по арифметическим правилам)

По усмотрению ведущего преподавателя, если это не противоречит действующим документам на момент экзамена, при получении накопленной оценки 8 баллов и более, студент может быть освобожден от экзамена. В таком случае, с согласия студента, ему выставляется результирующая оценка, равная накопленной.

Студент не получает возможность пересдать низкие результаты за домашнюю работу и/или контрольную работу, а также при пропуске соответствующих им учебных часов, при отсутствии уважительной причины пропуска.

При наличии уважительной причины пропуска (болезнь или другая документально подтвержденная причина), студент получает право пересдать пропущенные работы, при этом документы подтверждающие причину отсутствия, должны быть представлены в учебный офис, в течении недели со дня первого выхода на занятия после пропуска.

При получении неудовлетворительной оценки **О_{результ}** (значение после округления менее 4 баллов) выставляется оценка «НЕУДОВЛЕТВОРИТЕЛЬНО».

8 Образовательные технологии

Основными образовательными технологиями являются: интерактивные лекции, работа в группах на практических занятиях и семинарах.

9 Учебно-методическое и информационное обеспечение дисциплины

9.1 Основная литература

1. Murphy K. P. Machine learning: a probabilistic perspective [Electronic Resource] / Kevin P. Murphy.- Cambridge University Press, 2012. - 1098 p. - Authorized access: <http://site.ebrary.com/lib/hselibrary/detail.action?docID=10597102> (Online Digital Library "Ebrary").

9.2 Дополнительная литература

1. Hetland M. L. Python Algorithms: mastering basic algorithms in the Python Language [Electronic Resource] / Magnus Lie Hetland. – Apress, 2014. – 337 p. - Authorized access: <https://library.books24x7.com/toc.aspx?bookid=72529> (Online Digital Library "Books24x7").

10 Рекомендации для самостоятельной работы студентов.

Самостоятельная работа может рассматриваться как организационная форма обучения - система педагогических условий, обеспечивающих управление учебной деятельностью по освоению знаний и умений в области учебной деятельности без посторонней помощи. Студенту нужно четко понимать, что самостоятельная работа – не просто обязательное, а необходимое условие для получения знаний по дисциплине и развитию компетенций, необходимых в будущей профессиональной деятельности.

Самостоятельная работа проводится с целью:

- систематизации и закрепления полученных на лекциях теоретических знаний;
- углубления и расширения теоретических знаний;

- формирования умений использовать нормативную, правовую, справочную документацию и специальную литературу;
- развития познавательных способностей и активности студентов: творческой инициативы, самостоятельности, ответственности и организованности;
- формирования самостоятельности мышления, способностей к саморазвитию, самосовершенствованию и самореализации;
- формирования практических (общеучебных и профессиональных) умений и навыков;
- развития исследовательских умений;
- получения навыков эффективной самостоятельной профессиональной (практической и научно-теоретической) деятельности.

В учебном процессе выделяют два вида самостоятельной работы:

- аудиторная;
- внеаудиторная.

Аудиторная самостоятельная работа по дисциплине выполняется на учебных занятиях под непосредственным руководством преподавателя и по его заданию.

Внеаудиторная самостоятельная работа - планируемая учебная работа студентов, выполняемая во внеаудиторное время по заданию и при методическом руководстве преподавателя, но без его непосредственного участия.

Самостоятельная работа, не предусмотренная программой учебной дисциплины, раскрывающей и конкретизирующей ее содержание, осуществляется студентом инициативно, с целью реализации собственных учебных и научных интересов.

Для более эффективного выполнения самостоятельной работы по дисциплине преподаватель рекомендует источники для работы, характеризует наиболее рациональную методику самостоятельной работы, демонстрирует ранее выполненные студентами работы и т. п.

Виды заданий для внеаудиторной самостоятельной работы, их содержание и характер могут иметь вариативный и дифференцированный характер, учитывать индивидуальные особенности студента.

Самостоятельная работа может осуществляться индивидуально или группами студентов online и на занятиях в зависимости от цели, объема, конкретной тематики самостоятельной работы, уровня сложности.

Контроль результатов внеаудиторной самостоятельной работы осуществляется в пределах времени, отведенного на обязательные учебные занятия по дисциплине на практических занятиях.

Для представления результатов самостоятельной работы рекомендуется:

Составить план выступления, в котором отразить тему, самостоятельный характер проделанной работы, главные выводы и/или предложения, их краткое обоснование и практическое значение – с тем, чтобы в течение 3 – 5 минут представить достоинства выполненного самостоятельно задания.

Подготовить иллюстративный материал в виде презентации для использования во время представления результатов самостоятельной работы в аудитории. Конкретный вариант наглядного представления результатов определяется форматом аудиторного занятия и задания преподавателя.

В качестве дополнительной самостоятельной подготовки студентам рекомендуется использовать сайт <https://www.kaggle.com>, для участия в соревнованиях по построению моделей, а так же для изучения разбора лучших решений, для завершенных соревнований.



11 Материально-техническое обеспечение дисциплины и информационные технологии, используемые при осуществлении образовательного процесса по дисциплине, включая перечень программного обеспечения информационных справочных систем (при необходимости).

Для проведения семинарских занятий требуется компьютерный класс. Для проведения требуется аудитория, оснащенная проектором.

12 Особенности организации обучения для лиц с ограниченными возможностями здоровья

В случае необходимости, обучающимся из числа лиц с ограниченными возможностями здоровья (по заявлению обучающегося) могут предлагаться следующие варианты восприятия учебной информации с учетом их индивидуальных психофизических особенностей, в том числе с применением электронного обучения и дистанционных технологий:

1) *для лиц с нарушениями зрения:* в печатной форме увеличенным шрифтом; в форме электронного документа; в форме аудиофайла (перевод учебных материалов в аудиоформат); индивидуальные консультации с привлечением тифлосурдопереводчика; индивидуальные задания и консультации.

2) *для лиц с нарушениями слуха:* в печатной форме; в форме электронного документа; видеоматериалы с субтитрами; индивидуальные консультации с привлечением сурдопереводчика; индивидуальные задания и консультации.

3) *для лиц с нарушениями опорно-двигательного аппарата:* в печатной форме; в форме электронного документа; в форме аудиофайла; индивидуальные задания и консультации.