**Санкт-Петербургский государственный университет**

**Р А Б О Ч А Я П Р О Г Р А М М А**

**УЧЕБНОЙ ДИСЦИПЛИНЫ**

Машинное обучение

Machine Learning

**Язык(и) обучения**

русский

Трудоемкость в зачетных единицах: 5

Регистрационный номер рабочей программы: 051793

**Раздел 1. Характеристики учебных занятий**

**1.1. Цели и задачи учебных занятий**

Обучение известным алгоритмам в области машинного обучения.

**1.2. Требования подготовленности обучающегося к освоению содержания учебных занятий (пререквизиты)**

Аттестация по учебным дисциплина «Основы Python и SQL», «Программирование на С++», «Алгоритмы и структуры данных»

**1.3. Перечень результатов обучения (learning outcomes)**

Способность подбирать и программно реализовывать известные алгоритмы машинного обучения.

**1.4. Перечень и объём активных и интерактивных форм учебных занятий**

Практические занятия.

**Раздел 2. Организация, структура и содержание учебных занятий**

**2.1. Организация учебных занятий**

**2.1.1 Основной курс**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Трудоёмкость, объёмы учебной работы и наполняемость групп обучающихся | | | | | | | | | | | | | | | | | | |
| Код модуля в составе дисциплины,  практики и т.п. | Контактная работа обучающихся с преподавателем | | | | | | | | | | | | Самостоятельная работа | | | | Объём активных и интерактивных  форм учебных занятий | Трудоёмкость |
| лекции | семинары | консультации | практические  занятия | лабораторные работы | контрольные работы | коллоквиумы | текущий контроль | промежуточная  аттестация | итоговая аттестация | под руководством преподавателя | в присутствии  преподавателя | сам. раб. с использованием  методических материалов | текущий контроль (сам.раб.) | промежуточная аттестация (сам.раб.) | итоговая аттестация  (сам.раб.) |
| ОСНОВНАЯ ТРАЕКТОРИЯ | | | | | | | | | | | | | | | | | | |
| Форма обучения: очная | | | | | | | | | | | | | | | | | | |
| Семестр 6 | 30 |  |  | 15 |  |  |  |  | 2 |  |  |  | 19 |  | 6 |  | 20 | 2 |
|  | 2-100 |  |  | 10-25 |  |  |  |  | 10-25 |  |  |  | 1-1 |  | 1-1 |  |  |  |
| Семестр 7 | 30 |  | 2 | 30 |  |  |  |  | 2 |  |  |  | 26 |  | 18 |  | 20 | 3 |
|  | 2-100 |  | 2-100 | 10-25 |  |  |  |  | 2-100 |  |  |  | 1-1 |  | 1-1 |  |  |  |
| ИТОГО | 60 |  | 2 | 45 |  |  |  |  | 4 |  |  |  | 45 |  | 24 |  |  | 5 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Виды, формы и сроки текущего контроля успеваемости и промежуточной аттестации | | | | | | |
| Код модуля в составе дисциплины, практики и т.п. | Формы текущего контроля успеваемости | | Виды промежуточной аттестации | | Виды итоговой аттестации  (только для программ итоговой аттестации и дополнительных образовательных программ) | |
| Формы | Сроки | Виды | Сроки | Виды | Сроки |
| ОСНОВНАЯ ТРАЕКТОРИЯ | | | | | | |
| Форма обучения: очная | | | | | | |
| Семестр 6 |  |  | зачёт, устно, традиционная форма | по графику промежуточной аттестации |  |  |
| Семестр 7 |  |  | экзамен, устно, традиционная форма | по графику промежуточной аттестации |  |  |

**2.2. Структура и содержание учебных занятий**

Период обучения (модуль): **Семестр 6**

|  |  |  |  |
| --- | --- | --- | --- |
| **№ п/п** | **Наименование темы (раздела, части)** | **Вид учебных занятий** | **Кол-во часов** |
| 1 | **Байесовские сети**  1.1 Наивный Байесовский классификатор  1.2 Скрытый Байесовский классификатор  1.3 Скрытая Марковская модель  1.4 Алгоритм прямого-обратного хода  1.5 Алгоритм Витерби  1.6 Алгоритм Баума-Велша  1.7 Averaged one-dependence estimators (AODE) | лекции | 6 |
| практические занятия | 2 |
| сам. раб. с исп. метод. материалов | 3 |
| 2 | **Линейные классификаторы и их обобщения**  2.1 Метод опорных векторов  2.2 Линейная регрессия  2.3 Логистическая регрессия  2.4 Алгоритм Winnow  2.5 Квадратичный классификатор  2.6 Кернелизация (Kernel methods) | лекции | 6 |
| практические занятия | 4 |
| сам. раб. с исп. метод. материалов | 4 |
| 3 | **Классификация текстов**  3.1 Nearest centroid classifier  3.2 Классификатор Rocchio  3.3 Алгоритм Rocchio  3.4 Latent semantic analysis | лекции | 6 |
| практические занятия | 4 |
| сам. раб. с исп. метод. материалов | 4 |
| 4 | **Метод «k» ближайших соседей**  4.1 Метод «k» ближайших соседей (k-NN)  4.2 Large margin nearest neighbor | лекции | 6 |
| практические занятия | 2 |
| сам. раб. с исп. метод. материалов | 4 |
| 5 | **Выявление особенностей (Feature extraction)**  5.1 Метод главных компонент (PCA)  5.2 Kernel PCA  5.3 Многомерное шкалирование  5.4 Многофакторное понижение размерности (MDR)  5.5 Линейный дискриминантный анализ (LDA)  5.6 Isomap  5.7 Специальные методы для обработки изображений | лекции | 6 |
| практические занятия | 3 |
| сам. раб. с исп. метод. материалов | 4 |
|  | Промежуточная аттестация | самостоят. работа | 62 |
| зачет | 2 |
| **Итого** | | | **72** |

Период обучения (модуль): **Семестр 7**

|  |  |  |  |
| --- | --- | --- | --- |
| **№ п/п** | **Наименование темы (раздела, части)** | **Вид учебных занятий** | **Кол-во часов** |
| 6 | **Искусственные нейронные сети**  6.1 Персептрон  6.2 Метод обратного распространения ошибки  6.3 Автоэнкодер  6.4 Самоорганизующаяся карта Кохонена  6.5 Сеть Хопфилда  6.6 Сеть Хэмминга  6.7 Машина Больцмана  6.8 Ограниченная машина Больцмана  6.9 Сверточная нейронная сеть  6.10 Сиамская нейронные сети | лекции | 6 |
| практические занятия | 6 |
| сам. раб. с исп. метод. материалов | 4 |
| 7 | **Кластерный анализ**  7.1 Метод k-средних  7.2 Алгоритм HCS кластеризации  7.3 Алгоритм DBSCAN  7.4 Алгоритм OPTICS  7.5 Бикластеризация  7.6 Метод нечёткой кластеризации C-средних  7.7 Single-linkage clustering  7.8 Conceptual clustering (COBWEB) | лекции | 6 |
| практические занятия | 6 |
| сам. раб. с исп. метод. материалов | 5 |
| 8 | **Дерево принятия решений**  8.1 Алгоритм ID3  8.2 Алгоритм С4.5  8.3 Методы CART  8.4 Алгоритм VFDT  8.5 Случайный лес (Random forest) | лекции | 6 |
| практические занятия | 6 |
| сам. раб. с исп. метод. материалов | 5 |
| 9 | **Методы ассоциативного обучения**  9.1 Алгоритм Apriori  9.2 Алгоритм Eclat  9.3 Алгоритм FP-роста | лекции | 6 |
| практические занятия | 6 |
| сам. раб. с исп. метод. материалов | 6 |
| 10 | **Обучение с подкреплением**  10.1 SARSA  10.2 Q-обучение  10.3 Temporal difference learning | лекции | 6 |
| практические занятия | 6 |
| сам. раб. с исп. метод. материалов | 6 |
|  | Промежуточная аттестация | консультации | 2 |
| самостоят. работа | 182 |
| экзамен | 2 |
| **Итого** | | | **108** |

**Раздел 3. Обеспечение учебных занятий**

**3.1. Методическое обеспечение**

**3.1.1 Методические указания по освоению дисциплины**

Посещение лекционных занятий, выполнение самостоятельных работ.

**3.1.2 Методическое обеспечение самостоятельной работы**

Основная и дополнительная литература.

**3.1.3 Методика проведения текущего контроля успеваемости и промежуточной аттестации и критерии оценивания**

Методика проведения зачёта (6 семестр).

По дисциплине «Машинное обучение» в течение учебного года проводятся контрольные работы в соответствии с графиком и зачёт.

Обучающийся, не отчитавшийся по контрольной работе (контрольному домашнему заданию) в течение семестра, отчитывается по ним во время зачета. Для получения зачёта, кроме ответа на теоретические вопросы, необходимо решить задачу.

Критерии получения «зачёт»:

«Зачёт» выставляется, если выполняются три условия:

1. Экзаменуемым даны полные исчерпывающие ответы по всем вопросам, обучающийся свободно ориентируется в материале;

2. Экзаменуемый отвечает на все дополнительные вопросы

3. Экзаменуемый решил выданную задачу.

«Незачёт» выставляется, если не выполняются условия для получения «зачтено».

Оценка «зачтено» (**A**) ставится за полные ответы на вопросы билета, а также на дополнительные вопросы по теоретическому курсу, при условии, что в процессе всего ответа не было допущено грубых неточностей. за подготовку двух докладов-презентаций на выбранную обучающимся и согласованную с преподавателем тему; выступление с докладом на семинаре; грамотные ответы на вопросы преподавателя и слушателей по теме сообщения. А также за активную работу на семинарах: участие в обсуждении докладов других студентов, правильные ответы на вопросы преподавателя по другим темам семинара

Оценка «зачтено» (**B**) ставится за грамотные ответы на вопросы билета, возможно, с некоторыми неточностями, а также на дополнительных вопросы по курсу, которые задаются устно, обучающийся отвечает на них письменно и затем комментирует ответ устно. за подготовку двух докладов-презентаций на выбранную обучающимся и согласованную с преподавателем тему; выступление с докладом на семинаре; грамотные ответы на вопросы преподавателя и слушателей по теме сообщения.

Оценка «зачтено» (**С**) ставится за грамотные ответы на вопросы билета, возможно, с несущественными ошибками, а также на дополнительных вопросы по курсу, которые задаются устно, обучающийся отвечает на них письменно и затем комментирует ответ устно. за подготовку двух докладов-презентаций на выбранную обучающимся и согласованную с преподавателем тему; выступление с докладом на семинаре; грамотные ответы на вопросы преподавателя и слушателей по теме сообщения.

Оценка «зачтено» (**D**) ставится за содержательный в целом, но не полный письменный ответ на вопросы билета и на дополнительные вопросы по билету, без грубых ошибок,. выступление с 2 докладами на семинаре; грамотные ответы на вопросы преподавателя и слушателей по теме сообщения.

Оценка «зачтено» (**E**) ставится за содержательный в целом, но не полный письменный ответ на вопросы билета и на дополнительные вопросы по билету, без грубых ошибок, выступление с 2 докладами на семинаре; грамотные ответы на вопросы преподавателя и слушателей по теме сообщения.

Оценка «незачтено» (**F**) ставится в остальных случаях.

Методика проведения экзамена.

Экзамен проводится в устной форме. Билет состоит из двух вопросов. Время подготовки ответа на вопросы билета составляет 60 минут.

Использование конспектов и учебников, а также электронных устройств хранения, обработки или передачи информации при подготовке и ответе на вопросы экзамена не разрешается. В случае обнаружения факта использования недозволенных материалов (устройств) составляется акт, и студент удаляется с экзамена.

После ответа на вопросы билета преподаватель задает несколько дополнительных вопросов, на основании оценки ответов на которые итоговая оценка по предмету может быть повышена или понижена.

Критерии выставления оценок:

Оценка «отлично» ставится за полностью раскрытый теоретический материал и правильные ответы на дополнительные вопросы преподавателя.

Оценка «хорошо» ставится за изложенный теоретический материал билета (возможно с помощью наводящих подсказок преподавателя) и правильные ответы на дополнительные вопросы преподавателя.

Оценка «удовлетворительно» ставится за знание ответов на основные вопросы по каждой теме.

Оценка «неудовлетворительно» выставляется, если не выполняются условия для получения оценок «отлично», «хорошо» и «удовлетворительно».

Оценка «отлично» (A) ставится за полностью раскрытый теоретический материал и правильные ответы на дополнительные вопросы преподавателя.

Оценка «хорошо» (B) ставится за изложенный теоретический материал билета (возможно с помощью наводящих подсказок преподавателя) и правильные ответы на дополнительные вопросы преподавателя.

Оценка «хорошо» (C) ставится за полный ответ на вопросы билета с небольшими неточностями.

Оценка «удовлетворительно» (D) ставится за знание ответов на основные вопросы по каждой теме.

Оценка «удовлетворительно» (E) ставится за знание основных формулировок (по одной из каждого вопроса) с подсказками со стороны преподавателя.

Оценка «неудовлетворительно» (F) выставляется, если не выполняются условия для получения оценок «отлично», «хорошо» и «удовлетворительно».

**3.1.4 Методические материалы для проведения текущего контроля успеваемости и промежуточной аттестации (контрольно-измерительные материалы, оценочные средства)**

Примерный перечень вопросов к экзамену совпадает с темами обучения (см. раздел 2.2).

**3.1.5 Методические материалы для оценки обучающимися содержания и качества учебного процесса**

Оценка обучающимися содержания и качества учебного процесса по дисциплине осуществляется в установленном в СПбГУ порядке.

**3.2. Кадровое обеспечение**

**3.2.1 Образование и (или) квалификация штатных преподавателей и иных лиц, допущенных к проведению учебных занятий**

Базовое образование и/или ученая степень, соответствующая профилю дисциплины.

**3.2.2 Обеспечение учебно-вспомогательным и (или) иным персоналом**

Не требуется.

**3.3. Материально-техническое обеспечение**

**3.3.1 Характеристики аудиторий (помещений, мест) для проведения занятий**

Лекционная аудитория.

**3.3.2 Характеристики аудиторного оборудования, в том числе неспециализированного компьютерного оборудования и программного обеспечения общего пользования**

Мультимедийный проектор. Маркерная доска.

**3.3.3 Характеристики специализированного оборудования**

Не требуется.

**3.3.4 Характеристики специализированного программного обеспечения**

Не требуется.

**3.3.5 Перечень и объёмы требуемых расходных материалов**

Канцелярские принадлежности. Фломастеры цветные (из расчета 4 фломастера на 4 часа занятий). Губка для маркерной доски.

**3.4. Информационное обеспечение**

**3.4.1 Список обязательной литературы**

1. Гелиг А.Х., А.С. Матвеев. Введение в математическую теорию обучаемых распознающих систем и нейронных сетей. – СПб: Изд-во СПбГУ, 2014.

2. Rogers S., Girolami M. A First Course in Machine Learning. - CRC, 2017.- ЭР открытого доступа в сети интернет: http://ebook-dl.com/book/34346

3. Morhi M., Rostamizadeh A., Talwalkar A. Foundations of Machine Learning. - MIT, 2012. - ЭР открытого доступа в сети интернет

4. Alpaydin. E. Introduction to Machine Learning. - MIT, 2014. - ЭР открытого доступа в сети интернет: https://www.twirpx.com/file/1536456/

5. Kung. S.Y. Kernel Methods and Machine Learning. - MIT, 2014. - ЭР открытого доступа в сети интернет: https://www.twirpx.com/file/2045411/

6. Murphy Kevin. Machine Learning: A Probabilistic Perspective, MIT Press, 2012. – ЭР по подписке СПбГУ:

<https://ebookcentral.proquest.com/lib/stpeterst/detail.action?docID=3339490>.

7. Haykin S.. Neural Networks and Learning Machines. - Pearson, 2008. - ЭР открытого доступа в сети интернет:

<https://cours.etsmtl.ca/sys843/REFS/Books/ebook_Haykin09.pdf>.

8. Handbook of cluster analysis. - CRC, 2016. ЭР открытого доступа в сети интернет: https://www.twirpx.com/file/2165539/.

**3.4.2 Список дополнительной литературы**

1. Rasmussen K.E., Williams C.K.I. Gaussian Processes for Machine Learning. MIT-Cambridge, 2006.

2. Theodorides S. Machine Learning. A Bayesian and Optimization Perspective. Elsevier, 2015.

3. Everitt B.S., Landau S., Leese M., Stahl D. Cluster Analysis. Willey, 2011.

4. Handbook of Natural Language Processing. CRC, 2010.

**3.4.3 Перечень иных информационных источников**

Не предусмотрено.

**Раздел 4. Разработчики программы**

Ананьевский Михаил Сергеевич, доцент каф. Теоретической кибернетики,  
к.ф.-м.н. m.s.ananyevskiy@spbu.ru.