

Министерство образования и науки Российской Федерации  
Федеральное государственное автономное образовательное учреждение  
высшего образования «Новосибирский национальный исследовательский  
государственный университет» (Новосибирский государственный университет, НГУ)

**Механико-математический факультет**

УТВЕРЖДАЮ

Декан ММФ НГУ

\_\_\_\_\_ И.В. Марчук

«\_\_\_» \_\_\_\_\_ 201\_\_ г.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

**«Аналитика больших данных / Big Data Analytics» (спецсеминар)**

Направление подготовки: 01.04.02 ПРИКЛАДНАЯ МАТЕМАТИКА И ИНФОРМАТИКА

Профиль: Аналитика больших массивов данных /Big Data Analytics

Форма обучения: очная

Язык обучения: английский

Год обучения: 1 (2), семестр: 1-2 (3-4).

№	Вид деятельности	Семестр	Семестр
		1 (3)	2 (4)
1	Лекции, час.		
2	Практические занятия, час.	36	32
3	Лабораторные занятия, час		
4	Всего занятий в контактной форме, без учета промежуточной аттестации, час, из них		
5	в электронной форме, час.		
6	из них аудиторных занятий, час.		
7	из них в активной и интерактивной форме, час.		
8	консультаций, час.		
9	Самостоятельная работа, час.	34	38
10	в том числе на выполнение письменных работ, час		
11	Форма аттестации (экзамен, зачет, дифференцированный зачет), час	3, 2	3, 2
12	Всего зачетных единиц <sup>1</sup>	2	2

<sup>1</sup> С учетом выделенных часов на промежуточную аттестацию

Рабочая программа составлена на основании федерального государственного образовательного стандарта высшего образования по направлению подготовки «01.04.02 ПРИКЛАДНАЯ МАТЕМАТИКА И ИНФОРМАТИКА», утвержденного приказом Минобрнауки РФ от 28.08.2015 № 911.

Место дисциплины в структуре учебного плана: Блок «Дисциплины (модули)», вариативная часть, обязательные дисциплины.

Рабочая программа утверждена решением ученого совета Механико-математического факультета, протокол №\_\_ от \_\_.\_\_.\_\_\_\_г.

Программу разработал:

*Доцент* кафедры вычислительных систем ММФ,  
*К.ф.-м.н.*

Павловский Е.Н.

Согласовано:

Заведующий кафедрой вычислительных систем ММФ,

Глинский Б.М.  
Д.т.н.

## **Аннотация к рабочей программе дисциплины «Аналитика больших данных / Big Data Analytics» (спецсеминар)**

Семинар «Аналитика больших данных / Big Data Analytics» реализуется в рамках основной профессиональной образовательной программы (ОПОП) высшего образования по направлению подготовки «01.04.02 Прикладная математика и информатика» (очная форма обучения, язык реализации программы – английский), профиль – «Аналитика больших массивов данных». Он входит в вариативную часть, обязательные дисциплины блока «Дисциплины (модули)» образовательной программы и реализуется кафедрой вычислительных систем ММФ. Семинар проводится каждый учебный год, студенты обязаны посещать его как в первый год обучения (1-2 семестры), так и во второй (3-4 семестры).

Дисциплина направлена на формирование следующих компетенций:

ОПК-1: готовность к коммуникации в устной и письменной формах на государственном языке Российской Федерации и иностранном языке для решения задач профессиональной деятельности; в части следующих результатов обучения:

- ОПК-1.1 – владеть навыками создания на русском и/или иностранном языке письменных и устных текстов в области профессиональной деятельности;
- ОПК-1.2 – владеть навыками ведения научной дискуссии;

ОПК-3: способность самостоятельно приобретать с помощью информационных технологий и использовать в практической деятельности новые знания и умения, в том числе, в новых областях знаний, непосредственно не связанных со сферой деятельности, расширять и углублять свое научное мировоззрение; в части следующих результатов обучения:

- ОПК-3.1 – знать актуальные результаты и методы исследования в профессиональной области;

### **Содержание дисциплины:**

В рамках семинара студенты, специализирующиеся на кафедре вычислительных систем и кафедре программирования, и специалисты в области машинного обучения, больших данных из числа сотрудников кафедры, профильных лабораторий НГУ и других промышленных и научных организаций представляют доклады по материалам собственных исследований и актуальным результатам российских и зарубежных ученых.

Тематика семинара включает в себя широкий спектр задач, связанных исследованием моделей и методов машинного обучения, аналитики больших данных, построением и анализом точных и приближенных алгоритмов решения, в том числе:

- задачи компьютерного зрения,
- задачи обработки естественных языков,
- задачи распознавания и генерации речи,
- задачи анализа трафика.
- задачи информационного поиска,
- задачи кластеризации и классификации,
- задачи глубокого обучения
- и т.п.

Преподавание дисциплины предусматривает следующие виды учебной работы: практические занятия (участие в работе семинара, в том числе, представление собственных результатов и рефератов статей по тематике семинара). Самостоятельная

работа включает: разбор публикаций по тематике семинара, подготовку к выполнению доклада.

Общая трудоемкость дисциплины составляет 4 зачетных единицы.

#### **Правила аттестации по дисциплине.**

Текущий контроль формирования результатов обучения осуществляется в следующих формах:

- учет посещаемости занятий студентами, их активности в обсуждении докладов, участия в дискуссии;
- представления студентами докладов по материалам собственной научной работы или реферативного выступления с известными результатами по тематике семинара.

Промежуточная аттестация проводится в форме зачета в конце каждого семестра (отметка «зачет» или «незачет»). Решение о выставлении отметки принимается на основании результатов текущего контроля в ходе семестра.

#### **Учебно-методическое обеспечение дисциплины.**

1. Методические рекомендации по подготовке к докладу, курсовой и выпускной квалификационной работе, размещенные на сайте ММФ:  
<https://www.nsu.ru/n/mathematics-mechanics-department/studentam/thesis/>;
2. Перечень тем докладов семинара, размещенный на сайте  
[https://redmine.bigdata.nsu.ru/projects/scientific-workshop-big-data-analytics/wiki/Seminars\\_schedule](https://redmine.bigdata.nsu.ru/projects/scientific-workshop-big-data-analytics/wiki/Seminars_schedule).

## 1. Цели освоения дисциплины

Целью проведения семинара является ознакомление студентов с актуальными научными результатами в области машинного обучения и аналитики больших данных, выработка у студентов навыков и компетенций необходимых при проведении научно-исследовательской деятельности, представления результатов собственных научных изысканий.

## 2. Место дисциплины в структуре ООП

Семинар «Аналитика больших данных / Big Data Analytics» (далее – семинар) реализуется в рамках основной профессиональной образовательной программы (ОПОП) высшего образования по направлению подготовки «01.04.02 Прикладная математика и информатика» (очная форма обучения, язык реализации программы – английский), профиль – «Аналитика больших массивов данных» (2016-2017, 2017-2018 уч. годы) / «Анализ больших данных и искусственный интеллект» (2018-2019 уч. год). Он входит в вариативную часть блока «Дисциплины (модули)» образовательной программы и реализуется кафедрой вычислительных систем ММФ. Семинар проводится каждый учебный год, студенты обязаны посещать его как в первый год обучения (1-2 семестры), так и во второй (3-4 семестры).

## 3. Результаты освоения дисциплины

Семинар направлен на формирование следующих компетенций:

ОПК-1: готовность к коммуникации в устной и письменной формах на государственном языке Российской Федерации и иностранном языке для решения задач профессиональной деятельности; в части следующих результатов обучения:

- ОПК-1.1 – владеть навыками создания на русском и/или иностранном языке письменных и устных текстов в области профессиональной деятельности;
- ОПК-1.2 – владеть навыками ведения научной дискуссии;

ОПК-3: способность самостоятельно приобретать с помощью информационных технологий и использовать в практической деятельности новые знания и умения, в том числе, в новых областях знаний, непосредственно не связанных со сферой деятельности, расширять и углублять свое научное мировоззрение; в части следующих результатов обучения:

- ОПК-3.1 – знать актуальные результаты и методы исследования в профессиональной области;

Результаты изучения дисциплины	Формы организации занятий	
	Практич. занятия	Самост. Работа
ОПК-1.1 – владеть навыками создания на русском и/или иностранном языке письменных и устных текстов в области профессиональной деятельности	+	+
ОПК-1.2 – владеть навыками ведения научной дискуссии	+	+
ОПК-3.1 – знать актуальные результаты и методы исследования в профессиональной области	+	+

## 4. Структура и содержание дисциплины

В рамках семинара студенты, специализирующиеся на кафедре вычислительных систем и кафедре программирования, и специалисты в области машинного обучения, больших данных из числа сотрудников кафедры, профильных лабораторий НГУ и других индустри-

альных и научных организаций представляют доклады по материалам собственных исследований и актуальным результатам российских и зарубежных ученых.

Тематика семинара включает в себя широкий спектр задач, связанных исследованием моделей и методов машинного обучения, аналитики больших данных, построением и анализом точных и приближенных алгоритмов решения, в том числе:

- задачи компьютерного зрения,
- задачи обработки естественных языков,
- задачи распознавания и генерации речи,
- задачи анализа трафика.
- задачи информационного поиска,
- задачи кластеризации и классификации,
- задачи глубокого обучения
- и т.п.

Виды контактной работы	Лекции, час.	Практ., час.	Ссылки на результаты обучения
<b>Семестр 1 (3)</b>			
Участие в работе семинара, в том числе, представление собственных результатов и рефератов статей по тематике семинара.		36	ОПК-1.1 ОПК-1.2 ОПК-3.1
<b>Семестр 2 (4)</b>			
Участие в работе семинара, в том числе, представление собственных результатов и рефератов статей по тематике семинара.		32	ОПК-1.1 ОПК-1.2 ОПК-3.1

Участие в семинаре предполагает самостоятельную работу студентов.

№	Виды самостоятельной работы	Ссылки на результаты обучения	Часы на выполнение	Часы на консультации
<b>Семестр 1 (3)</b>				
1.	Самостоятельный разбор публикаций по тематике семинара	ОПК-1.1 ОПК-1.2 ОПК-3.1	12	
2.	Подготовка к докладу	ОПК-1.1 ОПК-1.2 ОПК-3.1	22	
<b>Семестр 2 (4)</b>				
3.	Самостоятельный разбор публикаций по тематике семинара	ОПК-1.1 ОПК-1.2 ОПК-3.1	16	
4.	Подготовка к докладу	ОПК-1.1 ОПК-1.2 ОПК-3.1	22	

### 5. Образовательные технологии

При реализации дисциплины применяются традиционные образовательные технологии. Занятия осуществляются в форме семинаров, в ходе которых проходит представление докладов и проводится научная дискуссия.

## 6. Учебно-методическое обеспечение самостоятельной работы студентов. Текущий контроль и промежуточная аттестация по учебной дисциплине

Текущий контроль формирования результатов обучения осуществляется в следующих формах:

- учет посещаемости занятий студентами, их активности в обсуждении докладов, участии в дискуссии;
- представления студентами докладов по материалам собственной научной работы или реферативного выступления с известными результатами по тематике семинара.

Промежуточная аттестация проводится в форме зачета в конце каждого семестра (отметка «зачет» или «незачет»). Решение о выставлении отметки принимается на основании результатов текущего контроля в ходе семестра.

В таблице ниже представлено соответствие форм аттестации заявляемым требованиям к результатам освоения дисциплины.

Результаты обучения	1 (3) семестр			2 (4) семестр		
	Участие в работе семинара	Доклад	Зачет	Участие в работе семинара	Доклад	Зачет
ОПК-1.1 – владеть навыками создания на русском и/или иностранном языке письменных и устных текстов в области профессиональной деятельности	+	+	+	+	+	+
ОПК-1.2 – владеть навыками ведения научной дискуссии	+	+	+	+	+	+
ОПК-3.1 – знать актуальные результаты и методы исследования в профессиональной области	+	+	+	+	+	+

Для обеспечения самостоятельной работы студентов используются следующие материалы:

- методические рекомендации по подготовке к докладу, курсовой и выпускной квалификационной работе, размещенные на сайте ММФ: <https://www.nsu.ru/n/mathematics-mechanics-department/studentam/thesis/>;
- Перечень тем докладов семинара, размещенный на сайте [https://redmine.bigdata.nsu.ru/projects/scientific-workshop-big-data-analytics/wiki/Seminars\\_schedule](https://redmine.bigdata.nsu.ru/projects/scientific-workshop-big-data-analytics/wiki/Seminars_schedule).

## 7. Литература

### (а) основная литература

1. Parija S. C., Kate V. (ed.). Writing and Publishing a Scientific Research Paper. – Springer, 2017. URL: <https://link.springer.com/content/pdf/10.1007%2F978-981-10-4720-6.pdf>. (Доступ из библиотеки НГУ: по базе Springer (п.9 разрешённых ресурсов НГУ: [http://libra.nsu.ru/dom\\_for\\_el\\_detabase/](http://libra.nsu.ru/dom_for_el_detabase/), SPRINGER (eBook Collection) - <http://link.springer.com/>).

### (б) дополнительная литература

Новиков, Александр Михайлович, Методология научного исследования: учебно-методическое пособие / А.М. Новиков, Д.А. Новиков; [науч. ред. Т.В. Новикова] Изд. Стер Москва : URSS: ЛИБРОКОМ, 2013, 270 с.

### (в) Интернет-ресурсы

1. Kiela D., Wang C., Cho K. Context-Attentive Embeddings for Improved Sentence Representations //arXiv preprint arXiv:1804.07983. – 2018. URL: <https://arxiv.org/abs/1804.07983>.

2. Kingma D. P. DP Kingma and M. Welling, Auto-encoding variational bayes, Proceedings of the 2nd International Conference on Learning Representations (ICLR), 2014, arXiv: 1312.6114 (2013). – 2013. URL: <https://arxiv.org/pdf/1312.6114.pdf>.

3. McInnes L., Healy J. Umap: Uniform manifold approximation and projection for dimension reduction //arXiv preprint arXiv:1802.03426. – 2018. URL: <http://arxiv.org/abs/1802.03426>.

4. Maaten L., Hinton G. Visualizing data using t-SNE //Journal of machine learning research. – 2008. – Т. 9. – №. Nov. – С. 2579-2605. <http://www.jmlr.org/papers/v9/vandermaaten08a.html>.

5. Frontiers in Massive Data Analysis [Electronic resource] / National Research Council // Washington, D.C.: The National Academies Press [2013]. – URL: <http://www.nap.edu/catalog/18374/frontiers-in-massive-data-analysis>. Read online (FREE): <https://www.nap.edu/read/18374>.

## 8. Учебно-методическое и программное обеспечение дисциплины

### 8.1. Учебно-методическое обеспечение

Методические рекомендации по подготовке к докладу, курсовой и выпускной квалификационной работе, размещенные на сайте ММФ: <https://www.nsu.ru/n/mathematics-mechanics-department/studentam/thesis/>; Перечень тем докладов семинара, размещенный на сайте [https://redmine.bigdata.nsu.ru/projects/scientific-workshop-big-data-analytics/wiki/Seminars\\_schedule](https://redmine.bigdata.nsu.ru/projects/scientific-workshop-big-data-analytics/wiki/Seminars_schedule).

### 8.2. Программное обеспечение

Для обеспечения реализации дисциплины используется стандартный комплект программного обеспечения (ПО), включающий регулярно обновляемое лицензионное ПО Windows, MS Office. Специализированное программное обеспечения для реализации курса не требуется.

ПО для лиц с ограниченными возможностями здоровья

№	Наименование ПО	Назначение	Место размещения
1	Jaws for Windows	Программа экранного доступа к системным и офисным приложениям, включая интернет-обозреватели. Информация с экрана считывается вслух, обеспечивая возможность речевого доступа к самому разнообразному контенту. Jaws также позволяет выводить информацию на обновляемый дисплей Брайля. JAWS включает большой набор клавиатурных команд, позволяющих воспроизвести действия, которые обычно выполняются только при помощи мыши.	Ресурсный центр, читальные залы библиотеки НГУ, компьютерные классы (сетевые лицензии)
2	Duxbury Braille	Программа перевода текста в текст	Ресурсный центр



	Translator v11.3 для Брайлевского принтера	Брайля, и печати на Брайлевском принтере	
3	"MAGic Pro 13" (увеличение+речь)	Программа для людей со слабым зрением и для незрячих людей. Программа позволяет увеличить изображение на экране до 36 крат, есть функция речевого сопровождения	Ресурсный центр, читальные залы библиотеки НГУ

### 9. Профессиональные базы данных и информационные справочные системы

1. Полнотекстовые журналы Springer Journals, электронные книги, реферативная БД по чистой и прикладной математике zbMATH.
2. Электронная библиотека диссертаций Российской государственной библиотеки (ЭБД РГБ)
3. Полнотекстовые электронные ресурсы Freedom Collection издательства Elsevier (Нидерланды) (коллекции Mathematics, Computer Science)
4. Электронные ресурсы Web of Science Core Collection (Thomson Reuters Scientific LLC.), Journal Citation Reports + ESI
5. Электронные БД JSTOR (США), предметная коллекция Mathematics & Statistics.
6. БД Scopus (Elsevier)
7. Лицензионные материалы на сайте eLibrary.ru

### 10. Материально-техническое обеспечение

#### Презентационное оборудование

№	Наименование	Назначение
1	Презентационное оборудование (мультимедиа-проектор, экран)	Для проведения лекционных занятий
2	Компьютерный класс (с выходом в Internet)	Для организации самостоятельной работы обучающихся

Оборудование, обеспечивающее адаптацию электронных и печатных образовательных ресурсов для обучающихся из числа лиц с ограниченными возможностями здоровья

№	Наименование оборудования	Назначение	Место размещения
1.	Принтер Брайля	Печать рельефно-точечным шрифтом Брайля	Ресурсный центр
2.	Увеличитель Prodigy Duo Tablet 24	Устройство для чтения и увеличения плоскочечного текста	Ресурсный центр, читальные залы библиотеки НГУ
3.	Специализированное мобильное рабочее место «ЭлНот 311»	Мобильный компьютер с дисплеем брайля	Ресурсный центр
4.	Портативный тактильный дисплей Брайля "Focus 40 Blue"	Навигация в операционных системах, программах и интернете с помощью отображения рельефно-точечным шрифтом Брайля получаемой информации	Ресурсный центр, читальные залы библиотеки НГУ
5.	Устройство для печати тактильной графики «PIAF»	Печать тактильных графических изображений	Ресурсный центр
6.	Портативный видеувеличитель RUBY XL HD	Увеличение текста и подбор контрастных схем изображения	Ресурсный центр

7.	Складной настольный электронный видеонадзор «ТОРАZ PHD 15»	Увеличение текста и подбор контрастных схем изображения	Ресурсный центр
8.	Электронный ручной видеонадзор ONYX Deskset HD 22”	Увеличение текста и подбор контрастных схем изображения	Ресурсный центр
9.	Смартфон EISmart G3	Смартфон клавишным управлением и озвученным интерфейсом, обучение спутниковой навигации.	Ресурсный центр
10.	FM-система «Сонет-PCM» PM-3-1	Звуковая FM-система для людей с нарушением слуха, улучшающая восприятие голосовой информации	Большая физическая аудитория главного корпуса НГУ