

МИНОБРНАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
«Новосибирский национальный исследовательский государственный университет»
(Новосибирский государственный университет, НГУ)
**Структурное подразделение Новосибирского государственного университета –
Специализированный учебно-научный центр Университета (СУНЦ НГУ)**
Министерство науки и высшего образования Российской Федерации

СОГЛАСОВАНО Заместитель директора по УР  (Петровская О.В.) 23 ноября 2023 г.	УТВЕРЖДЕНО На заседании ученого совета СУНЦ НГУ Протокол № 48 от 23 ноября 2023 г.	УТВЕРЖДАЮ Директор СУНЦ НГУ  (Некрасова Л.А.) 23 ноября 2023 г.
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------

РАБОЧАЯ ПРОГРАММА
курса внеурочной деятельности «Матлингвистика»

Заведующий лабораторией инженерного конструирования
Якушкин Сергей Владимирович



Новосибирск 2023

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Спецкурс рассчитан на учащихся 10-11 класса. Для посещения спецкурса требуются базовые школьные знания информатики, желателен интерес к лингвистике.

Автоматическая обработка естественного языка – это одно из направлений искусственного интеллекта и математической лингвистики. Задачи, связанные с обработкой естественного языка направлены на создание более удобной формы взаимодействия человека с электронными устройствами.

Цель курса – познакомить школьников с современными разработками в сфере автоматической обработки естественного языка.

ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ

Цели и задачи специального курса:

- 1) сформировать у учащихся знания о методах обработки текстов на естественном языке, развить умение работать с современными инструментами обработки текстов;
- 2) развить навыки программирования и разработки алгоритмов;
- 3) овладеть навыками постановки задачи и самостоятельной организации своей работы;
- 4) сформировать понимание принципов проектной деятельности и навыков презентации проведённой работы.

В результате освоения специального курса обучающийся должен:

- 1) научиться осуществлять самостоятельный поиск информации с использованием различных образовательных ресурсов (учебных текстов, официальной документации языка программирования, документации сторонних библиотек);
- 2) овладеть способностями к обобщению и анализу информации, постановке задачи и выбору путей ее решения;
- 3) уметь применять полученные знания и навыки на практике;
- 4) освоить программирование на языке Python и основные инструменты для обработки текстовой информации с помощью Python;
- 5) освоить понятия лингвистики, необходимые для работы с текстовой информацией.

В ходе спецкурса учащиеся должны работать индивидуально и в группах, разработать программу на основе инструментов автоматической обработки языка.

По результатам работы и по посещению будет выставлен зачёт. Зачёт недифференцированный (зачёт / незачёт).

Объём спецкурса – 48 академических часов.

Планируемое время проведения спецкурса – 2023-24 учебный год (сентябрь 2023 г. - февраль 2024 г.).

СОДЕРЖАНИЕ СПЕЦКУРСА

Занятие 1 (2 часа)

Алгоритмы. Программирование. Особенности языка Python. Переменные. Типы данных.

Занятие 2 (2 часа)

Функции print, input. Целочисленный тип данных. Арифметические и логические операции. Сравнение. Управляющие структуры.

Занятие 3 (2 часа)

Строковый тип данных. Операции со строками.

Занятие 4 (2 часа)

Методы и функции для работы со строками. Сравнение строк.

Занятие 5 (2 часа)

Условия. Структуры if, else, elif. Логические выражения с and, or, not.

Занятие 6 (2 часа)

Циклы. Описание и алгоритмы работы. Цикл while.

Занятие 7 (2 часа)

Цикл for. Функция range. Операторы break и continue.

Занятие 8 (2 часа)

Списки. Описание и принципы работы.

Занятие 9 (2 часа)

Кортежи. Отличия от списков. Операции и встроенные функции.

Занятие 10 (2 часа)

Словари. Принцип работы и операции со словарями.

Занятие 11 (2 часа)

Файлы. Расширения. Директории. Пути. Режимы работы с файлами.

Занятие 12 (2 часа)

Чтение и запись файлов в Python.

Занятие 13 (2 часа)

Пользовательские функции. Структура и особенности. Входные и выходные данные

Занятие 14 (2 часа)

Классы. Терминология. Работа с пользовательскими классами.

Занятие 15 (2 часа)

Модули. Библиотеки. Импорт библиотек.

Занятие 16 (2 часа)

Регулярные выражения. Методы и функции.

Занятие 17 (2 часа)

Морфологический анализ. Понятие. Библиотека rymorphy2. Морфологический анализ существительных и оценка вероятностей.

Занятие 18 (2 часа)

Теги. Атрибуты тегов. Нормальная форма слова. Парадигма и предсказание.

Занятие 19 (2 часа)

Набор библиотек NLTK. Обзор возможностей. Функции и методы.

Занятие 20 (2 часа)

Токенизация слов и предложений. Частеречная разметка и разметка по частям предложения.

Занятие 21 (2 часа)

Семантический анализ. Использование WordNet.

Занятие 22 (2 часа)

Графический интерфейс. Модуль TKInter. Методы, функции, события.

Занятие 23 (2 часа)

Основы объектно-ориентированного программирования. Инкапсуляция.

Занятие 24 (2 часа)

Полиморфизм, Наследование.

ТЕМАТИЧЕСКОЕ ПЛАНИРОВАНИЕ

№ п/п	Наименование разделов и тем программы	Кол-во часов	Воспитательный компонент
1	Алгоритмы. Программирование. Особенности языка Python. Переменные. Типы данных.	2	<p>Формирование ответственного отношения к своему здоровью и безопасности. Готовность обучающихся к саморазвитию, самостоятельности и личному самоопределению. Формирование способности к оценке своих возможностей. Наличие мотивации к целенаправленной социально значимой деятельности.</p>
2	Функции print, input. Целочисленный тип данных. Арифметические и логические операции. Сравнение. Управляющие структуры.	2	
3	Строковый тип данных. Операции со строками.	2	
4	Методы и функции для работы со строками. Сравнение строк.	2	
5	Условия. Структуры if, else, elif. Логические выражения с and, or, not.	2	
6	Циклы. Описание и алгоритмы работы. Цикл while.	2	
7	Цикл for. Функция range. Операторы break и continue.	2	
8	Списки. Описание и принципы работы.	2	
9	Кортежи. Отличия от списков. Операции и встроенные функции.	2	
10	Словари. Принцип работы и операции со словарями.	2	
11	Файлы. Расширения. Директории. Пути. Режимы работы с файлами.	2	
12	Чтение и запись файлов в Python.	2	
13	Пользовательские функции. Структура и особенности. Входные и выходные данные.	2	
14	Классы. Терминология. Работа с пользовательскими классами.	2	
15	Модули. Библиотеки. Импорт библиотек.	2	
16	Регулярные выражения. Методы и функции.	2	
17	Морфологический анализ. Понятие. Библиотека rymorphy2. Морфологический анализ сущностей и оценка вероятностей.	2	<p>Развитие и поддержка одаренности обучающихся и обеспечение участия в олимпиадах и конкурсах Установление доверительных отношений между руководителем объединений и обучающимися и между обучающимися непосредственно через беседы, дискуссии</p>
18	Теги. Атрибуты тегов. Нормальная форма слова. Парадигма и предсказание.	2	
19	Набор библиотек NLTK. Обзор возможностей. Функции и методы.	2	
20	Токенизация слов и предложений. Частеречная разметка и разметка по частям предложения.	2	
21	Семантический анализ. Использование WordNet.	2	
22	Графический интерфейс. Модуль TKInter. Методы, функции, события.	2	

23	Основы объектно-ориентированного программирования. Инкапсуляция.	2	
24	Полиморфизм, Наследование.	2	
	Всего	48	

Приложение 1

УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА

1. ИНТУИТ (Национальный открытый университет):
<http://www.intuit.ru/studies/courses/49/49/info> Открытый курс.
2. Natural Language Processing with Python. – O'Reilly Media, 2009.
3. Dan Jurafsky and James H. Martin. Speech and Language Processing. - Prentice-Hall, 2000
4. Марк Лутц. Изучаем Python. 4-е издание. Перевод А. Киселева. – Символ, 2011.
5. Mark Lutz. Learning Python. 5th edition. – O'Reilly Media, 2013.
6. Документация Pymorphy2 <http://pymorphy2.readthedocs.org/en/latest/user/>
7. Учебник по NLTK <http://www.nltk.org/book/>