



# Дипломная работа

профессия Data Scientist

курс Машинное обучение



**Вероника Лобанова**  
координатор  
направления Data Science



# План на сегодня

**18:00 - 18:40**

1. Документы для допуска
2. Выбор темы
3. Выбор ментора
4. Защита дипломов
5. Ответы на вопросы



# Документы



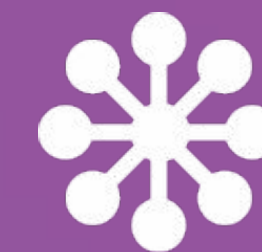
# — Студенты курса Машинное обучение

1. Паспорт (первая страница + прописка)
2. Диплом о высшем или среднем специальном образовании
3. Справка о смене фамилии (если менялась)



# — Студенты профессии Data Scientist

1. Копия паспорта
2. Копия документов об образовании
3. Справка о смене фамилии
4. Заявление о зачислении на курс
5. Договор



# Выбор темы

Выбираете тему сами исходя из критериев оценки

## КРИТЕРИИ ОЦЕНКИ



# **ПРИМЕРЫ ДИПЛОМНЫХ РАБОТ СТУДЕНТОВ**

---

# Временные ряды

1.

*Предсказание*  
трудозатрат, расходов  
по проектам *во*  
*времени*. Детекция  
аномалий и их  
прогнозирование.

2.

Автоматизированное  
*выявление патологии*  
на снимках ЭКГ

3.

*Прогнозирование цен*  
биржевых активов:  
акций, фьючерсов,  
валют в масштабе от  
секунд до дней

---

# Компьютерное зрение

1.

Распознавание

*аномалий*

*распределения кур* по

площади пола ангара

для отслеживания

здоровья кур

2.

Поиск и

*категоризация*

*предметов одежды*

на фотографии

3.

*Определение темы*

*книги по обложке* и

названию —

многоклассовая

мульти-label

классификация

---

# NLP

## 1.

Чат бот первой  
линии службы IT  
поддержки: *поиск  
ответов на запросы*  
по базе знаний

## 2.

*Обнаружение  
политических  
ботов* в  
социальных сетях

## 3.

*Оценка тональности и  
автоматический  
перевод* отзывов о  
мобильном  
приложении

---

# Немного бизнеса

1.

Прогноз *вероятности*  
*конверсии* лидов  
в CRM Битрикс24

2.

Прогноз продаж  
покупателя на три  
недели вперёд

3.

Модель склонности  
*клиента* к карте  
рассрочки

---

# Рекомендательные системы

1.

*Тут все понятно, кейсы  
довольно  
предсказуемые.*

2.

Последние 5 лет весь  
мир писал по этой  
теме дипломы.  
Поднадоело, да и  
нейронки сильно  
развились.

3.

*Но! Это весь онлайн-  
ритейл, телеком, банки  
и не только.*

# Примеры. Медицина

- 1/ Определение заболеваний по рентген-снимкам и МРТ (временные ряды + CV)
- 2/ Диагностика заболеваний кожи и глаз (CV)
- 3/ Прогнозирование лечения для зубов и суставов / отказа протезов, поиск оптимального решения (много разных вариантов, и с CV, и без)
- 4/ Анализ состояние посетителей и пациентов мед.учреждения при помощи внешних камер слежения — фактически, визуальная диагностика (CV)
- 5/ Рекомендательная система подбора лекарств и анализов по симптомам и/или истории болезни (рексис + NLP)

# Примеры. Медицина

- 6/ Предсказание эффектов генетической изменчивости, необходимой для интерпретации ДНК
- 7/ Создание новых препаратов и предсказание побочных эффектов применения (классический ML)
- 8/ Фссистент-графический интерфейс для пациентов и посетителей мед.учреждения (чатботы и рексис)
- 9/ Решение для data management электронных медкарт (частично с применением NLP)
- 10/ Сегментация изображений для микрохирургии или печати компонентов



# Примеры. Банки

- 1/ ИИ финансовый тренер / консультант (можно через чатбота, значит — NLP)
- 2/ ассистент для поиска транзакций в исторических данных + вычисление суммы кредита и долга — вместо ручных подсчетов в калькуляторе в моб.приложении
- 3/ бот для банковских сотрудников, который ищет инсайты в транзакциях клиентов и помогает на основании паттернов покупок разрабатывать новые кредитные продукты
- 4/ скоринговая система — это нестареющая классика (можно используя XGBoost или ANN)
- 5/ распознавание текстов договоров, форм, заявок, претензий — тоже классика, но все еще востребованная(OCR / NLP)

# Примеры. Банки

- 6/ предсказание оттока клиентов на основании конкурентного анализа банковских продуктов-конкурентов (классический ML)
- 7/ модель алгоритмического трейдинга — можете обратить внимание на то, что делает AimTech (российская компания, к слову)
- 8/ распознавание спутниковых изображений для предсказания частоты покупок в определенных магазинах, потоков поставок, маршрутов и т.д. (CV + классический ML)
- 9/ модель расчета стоимости актива, по аналогии с поведением сотрудника-оценщика + исторические данные по похожим активам — такое востребовано в инвест-банкинге

# Для вдохновения

- [Презентаци DS-3](#)
- [Примеры тем](#)



# Ментор

- При подготовке проекта вам будет помогать наставник — опытный специалист в предметной области или решении класса задач, к которому относится ваш проект. Его задача — дать обратную связь по работе и, если нужно, скорректировать направление, в котором ищете решение.
- Формат работы — онлайн-консультации голосом. Продолжительность — консультации по 30 минут в общей сложности до 4 часов.

## Памятка по работе с ментором



# Защита

# Защита диплома перед дипломной комиссией

- Отчет об исследовании + презентация + код
- 10 минут на выступление + 10 минут на вопросы  
КОМИССИИ



## **Даты защиты**

**DS-6 + AML-4 - 22 декабря**

**DS-7 + AML-5 - 15 марта**

**DS-8 + AML-6 - 8 августа**





**Q&A**



**Спасибо за внимание**

Вероника Лобанова

[v.lobanova@netology-group.ru](mailto:v.lobanova@netology-group.ru)