

Старостина Лилия Валерьевна

Телефон: +7 914 283 09 23
email: l.v.starostina2014@gmail.com
Telegram: @Lily_Val
Портфолио:
<https://starostinalv.github.io/>

Data Scientist

Стек:

SQL, Python, библиотеки: Numpy, Pandas, SciPy, Statsmodels, NLTK, Re, Pymorphy2, Matplotlib, Seaborn, Sklearn, Pymystem3, Genism, Pytorch, Keras, Cv2, PIL, Pyspark, Surprise, Requests, BeautifulSoup, Time

Ключевые компетенции:

- Работа с библиотеками Python для подготовки и анализа данных
- Написание SQL-запросов для получения данных
- Работа с большими данными
- Создание и обучение нейросетей
- Работа с текстовыми данными, изображениями и временными рядами
- Создание рекомендательных систем
- Работа с признаками, выбор и построение моделей классического машинного обучения

Профессиональный опыт

август 2022 г. – настоящее время

Учебные проекты в ООО «Нетология»

Направление «Аналитика», курс «Data Scientist» (программа курса [по ссылке](#))

SQL

Задачи: написание запросов к БД авиаперевозок. **Технологии:** PostgreSQL, DBeaver, работа с данными различных типов, соединения join, агрегатные функции, группировка и фильтрация, оконные функции, CTE, рекурсия, материализованные представления.

Обработка естественного языка

Задачи:

- классификация отзывов по тональности, тематическое моделирование отзывов;
- составление словарей для классификации текстов по тональности, распространение метки;
- перевод фраз с помощью механизма внимания.

Технологии: регулярные выражения, токенизация, лемматизация, эмбединги, word2vec, TF-IDF, LSI, «облако слов», countvectorizer, truncatedSVD, LDA, CNN, LSTM, TSNE, Label propagation, f1, модель «seq2seq with attention»

Компьютерное зрение

Задачи:

- распознавание рукописных символов. **Технологии:** гистограмма градиентов, KMeans;
- классификация изображений («Cats vs Dogs»). **Технологии:** VGG16, InceptionV3;
- сегментация объектов. **Технологии:** U-Net;
- детекция изображений клеток крови. **Технологии:** VGG, MultiBox, SSD;

- улучшение качества обучения нейросети для классификации изображений.
- Технологии:** EfficientNet, LR Schedulers, аугментация, сеть как Feature Extractor, Fine Tuning.

Big-Data

Задачи: BI-аналитика данных о фильмах, рекомендательная система фильмов.

Технологии: предсказание оценки фильма с помощью линейной регрессии средствами Pandas; подсчет средней оценки фильмов, поиск самых популярных и непопулярных фильмов средствами PySpark.

Анализ временных рядов

Задачи: SSA временного ряда; построение моделей MA, ARIMA, GARCH, оценка ряда через HMM. **Технологии:** SSA, матрица корреляции, EMA, TEMA, ARIMA, GARCH, матрица переходных вероятностей, случайные марковские процессы.

Рекомендательные системы

Задача: создание моделей для рекомендации фильмов. **Технологии:** TF-IDF, линейная регрессия, RMSE, коллаборативная фильтрация, гибридизация.

Построение моделей классического ML

Задачи: классификация, регрессия, кластеризация объектов. **Технологии:** логистическая и линейная регрессия, дерево решений, метод опорных векторов, кластеризация (KMeans, DBSCAN, AgglomerativeClustering), метрики TPR, FPR, ROC-кривая, Precision, Recall, RMSE, R2; подготовка сырых данных, EDA, Feature Selection, Feature Engineering, алгоритмы оптимизации параметров моделей.

июль 2019 г. – май 2022 г.

Сервисный координатор

ООО «Сандвик майнинг энд констракшн СНГ», г. Нерюнгри

- Собирала данные о работе сервисной службы, вела базы «Service Work Orders» и «Сервисный модуль»; подготавливала ежемесячные отчетные документы по работе сервисной службы;
- Собирала данные и подготавливала отчеты по БТ и ООС в системе «EHS-360»; занималась приведением документации подразделения в области БТ и ООС в соответствие с требованиями компании; устраняла замечания по ТБ;
- Занималась заведением регулярных счетов в систему Procurement, вела контроль оплаты.
- Осуществляла устные и письменные переводы (английский язык).

декабрь 2017 г. – июль 2018 г.

Инженер-аналитик, АО ХК «Якутуголь», г. Нерюнгри

- Разрабатывала и вела аналитическую отчетность;
- Анализировала работу горно-транспортного и вспомогательного оборудования на основе данных первичной фактуры, систем «MODULAR MINING SYSTEMS» и «AutoGRAPH».

август 2006 г. – январь 2016 г.

Работа в ГОУ ВПО ТИ(ф) ЯГУ (г. Нерюнгри) в должностях ассистента и старшего преподавателя, НГТУ (г. Новосибирск) в должности старшего преподавателя:

- Читала курсы дисциплин «ТАУ», «Метрология», «Материаловедение. ТКМ», «Электротехника и электроника» и др.;
- Разработала учебно-методические комплексы (УМК) и электронные УМК дисциплин;
- Руководила ВКР студентов;

- Участвовала в подготовке документов кафедры по самообследованию перед аккредитацией;
- Занималась научной работой (опубликовано 22 статьи, 1 монография).

Работа в МУ «Крытый стадион «Горняк» в должности инженера-энергетика:

- Выполняла обязанности ответственного за электрохозяйство, за пожарную безопасность и охрану труда;
- Подготовила программы и инструкции по ОТ, разработала планы пожарной эвакуации из помещений Стадиона;
- Проводила инструктажи, участвовала в аттестационных комиссиях по ОТ и ПБ.

Образование

2001 – 2006 гг. ГОУ ВПО «Якутский государственный университет им. М.К. Аммосова», Технический институт (филиал) в г. Нерюнгри, инженерный факультет. Диплом с отличием о высшем образовании, квалификация инженера по специальности «Электропривод и автоматика промышленных установок и технологических комплексов»

Дополнительное образование

- **2021 – 2024 гг.** ООО «Нетология», направление «Аналитика», курс «Data Scientist» (программа курса [по ссылке](#));
- **2023 г.** «Amazon Online Learning», курсы «Math for Machine Learning», «Linear and Logistic Regression», «Data Analytics Fundamentals»;
- **2008 – 2010 гг.** ГОУ ВПО «ЯГУ им. М.К. Аммосова». Диплом о дополнительном (к высшему) образовании. Дополнительная квалификация преподавателя высшей школы;
- **2002 – 2010 гг.** Повышение квалификации в области энергосбережения, педагогики, ИТ.

Дополнительная информация

- Иностранные языки: английский (B2 – Upper Intermediate) китайский (A1 – Beginner)
- Водительское удостоверение категории В
- Готова к командировкам и релокации
- Хобби: спорт, комнатные цветы, иностранные языки

Обо мне

Работала в образовательной и добывающей отраслях. Изучила курс «Data science» в ООО «Нетология». Имею опыт сбора и анализа данных о работе промышленного оборудования и данных для научных исследований в области энергетики, а также опыт написания научных статей. Решила продолжить профессиональное развитие в аналитике, т.к. считаю данное направление интересным, актуальным и перспективным.