



II ежегодный митап для вузов Loginom.Университеты
22 января 2026

Внедрение ИИ в подготовку аналитиков данных: опыт магистратуры СПбПУ

Николай Паклин

О докладчике



- Главный исследователь данных в Loginom, руководитель академической программы
- 20+ лет в low-code аналитике данных
- К.т.н., доцент СПбПУ Петра Великого, РЭУ имени Г.В. Плеханова, автор книг и учебников по аналитике данных

Благодарности

Работы осуществляются в рамках проекта
**«Аналитика самообслуживания и искусственный
интеллект: редизайн курса по машинному
обучению»**, реализуемого победителем грантового
конкурса для преподавателей 2024/2025
Стипендиальной программы Владимира Потанина от
СПбПУ.

Партнер проекта – Loginom.



страница проекта



Характеристика магистерской программы

Направление «Бизнес-информатика», программа «Прикладное машинное обучение в управлении бизнесом», ВШБИ СПбПУ

Цели программы: формирование у выпускников компетенций, обеспечивающих проектирование и внедрение интеллектуальных систем поддержки принятия решений в контексте цифровой трансформации бизнеса.

Методологическая основа: интеграция принципов машинного обучения (ML), аналитики данных и концепций управления бизнес-процессами (Business Process Management).

Ключевые компетенции:

- применение алгоритмов ML для решения прикладных бизнес-задач;
- рутинизация бизнес-процессов при помощи технологий ML и ИИ;
- интерпретация результатов анализа данных и представление их в формате, понятном для лиц, принимающих решения;
- оценка экономической эффективности внедрения ML-решений.



Вызовы подготовки специалистов по машинному обучению

- Рост объемов и разнообразия данных
- Дефицит ИТ-специалистов
- Гетерогенность входных компетенций магистрантов **(ключевая проблема)**
- Стремительное развитие ИИ **(нельзя игнорировать)**
- Недостаток аудиторных часов для ряда тем в учебном плане



*Входной опрос студентов 1-го курса магистратуры

Задачи проекта

- Выравнивание уровней компетентности студентов с разным опытом программирования
- Внедрение гибридной модели обучения: low-code и AI-coding
- Снижение барьеров входа в профессию аналитика данных
- Интеграция современных ИИ-инструментов в учебный процесс

Профиль магистрантов 1 курса

- 88 % используют ИИ-инструменты как пользователи
- 83 % не владеют опытом работы с базовыми библиотеками Pandas/NumPy в Python
- 71 % не знакомы с системой контроля версий Git
- 59 % заинтересованы в low-code инструментах

Результаты входного опроса 2025/26 года.

По предшествующим годам ситуация в целом похожая.

Системный подход: комплексное решение

Цель: обеспечить эффективное освоение аналитики данных и Python, независимо от начального уровня, за счет аналитики самообслуживания, сквозных примеров, методических указаний с акцентом на использовании Open Source решений и поддержке отечественных вендоров.

Компоненты подхода

Аналитика самообслуживания

- **Low-code:** разведочный анализ, освоение базовых операций с датасетами, WoE-анализ, построение ML-модели, метрики качества
- **LLM:** Помощь в Python-кодинге, код-ревью и оптимизация кода
- **AutoML:** Автоматизация выбора и оптимизации моделей, изучение ML-алгоритмов, улучшение метрик

Язык программирования Python

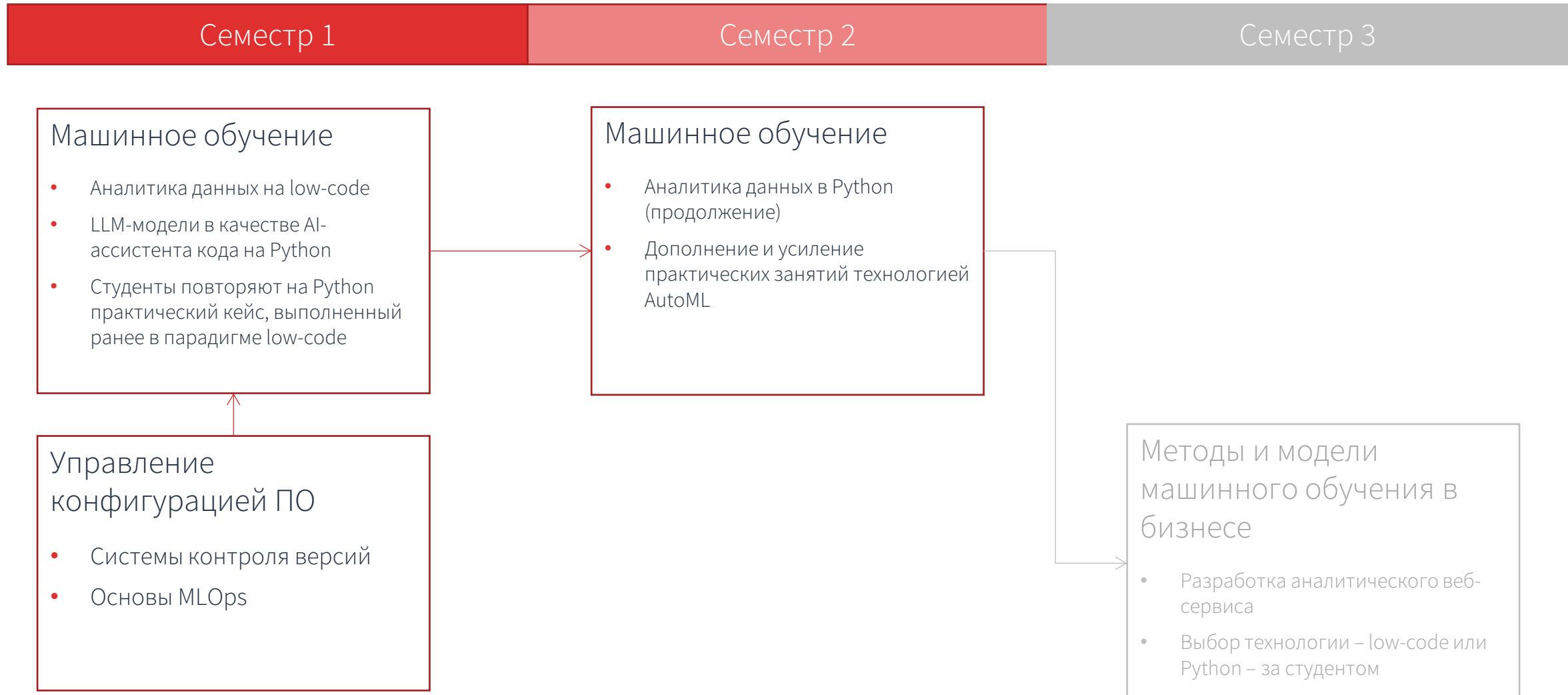
- Библиотеки **pandas** для манипуляции датасетами
- Библиотека **ydata_quality** для разведочного анализа
- Библиотека **optbinning**
- для WoE- и IV-анализа
- Библиотека **scikit-learn** для построения моделей и расчета метрик
- **Git** для контроля версий кода

Практический кейс

Датасет – транзакции розничной сети

- Выдвижение гипотез
- Подготовка выборок для моделирования
- Генерация признаков и анализ их значимости
- Модель вероятности повторного визита клиента

Реализация системного подхода в контексте учебных дисциплин



Технологический стек: используемое ПО

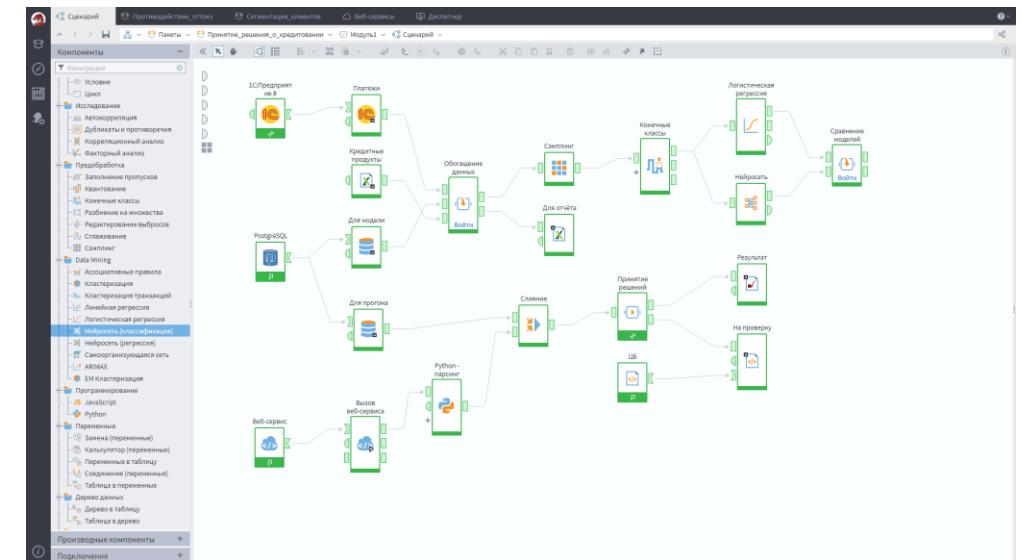
Компонент системного подхода	Программное обеспечение	Доступность	Отечественная разработка?	Краткое описание
Аналитика самообслуживания	Loginom	Бесплатно (для образования) / Платно	Да	Low-code аналитическая платформа с визуальным интерфейсом для подготовки, анализа и моделирования данных
	LightAutoML	Бесплатно (Open Source)	Да (Sber Tech)	Фреймворк автоматизированного машинного обучения для Python
	GigaCode, Continue.Dev	Бесплатно	Да (Sber Tech)	Плагины для среды VS Code
Язык программирования	Python	Бесплатно (Open Source)	Нет	Язык программирования для анализа данных и машинного обучения
Контроль версий	Git	Бесплатно (Open Source)	Нет	Система контроля версий для отслеживания изменений в коде
Платформа для совместной разработки	GitVerse	Бесплатно / Платно	Да (Sber Tech)	Платформа для хостинга Git-репозиториев, совместной разработки и управления проектами
Среда разработки	VS Code	Бесплатно	Нет	Среда разработки
MLOps и управление экспериментами	ClearML	Бесплатно/ Платно	Нет	Платформа MLOps. Есть облачный доступ с бесплатным тарифом, возможность установки в закрытом контуре

Loginom – low-code платформа, позволяющая бизнес-пользователям самостоятельно, (почти) без программирования, проектировать сложные аналитические процессы.

- Полностью российская разработка, в которой воплощен 25-летний опыт компании
- Входит в реестр отечественного ПО
- Импортозамещение ушедших вендоров
- Фокус на решении задачи, а не на синтаксисе языка программирования
- Невозможно сделать при помощи ИИ (пока что)



Loginom



GigaCode от Sber Tech

AI-ассистент кодинга для Python

Функциональность:

- автоматическая генерация Python-кода из пользовательских запросов;
- семантический анализ кода и объяснение синтаксиса;
- интеллектуальная поддержка отладки и исправление ошибок.

Преимущества для обучения:

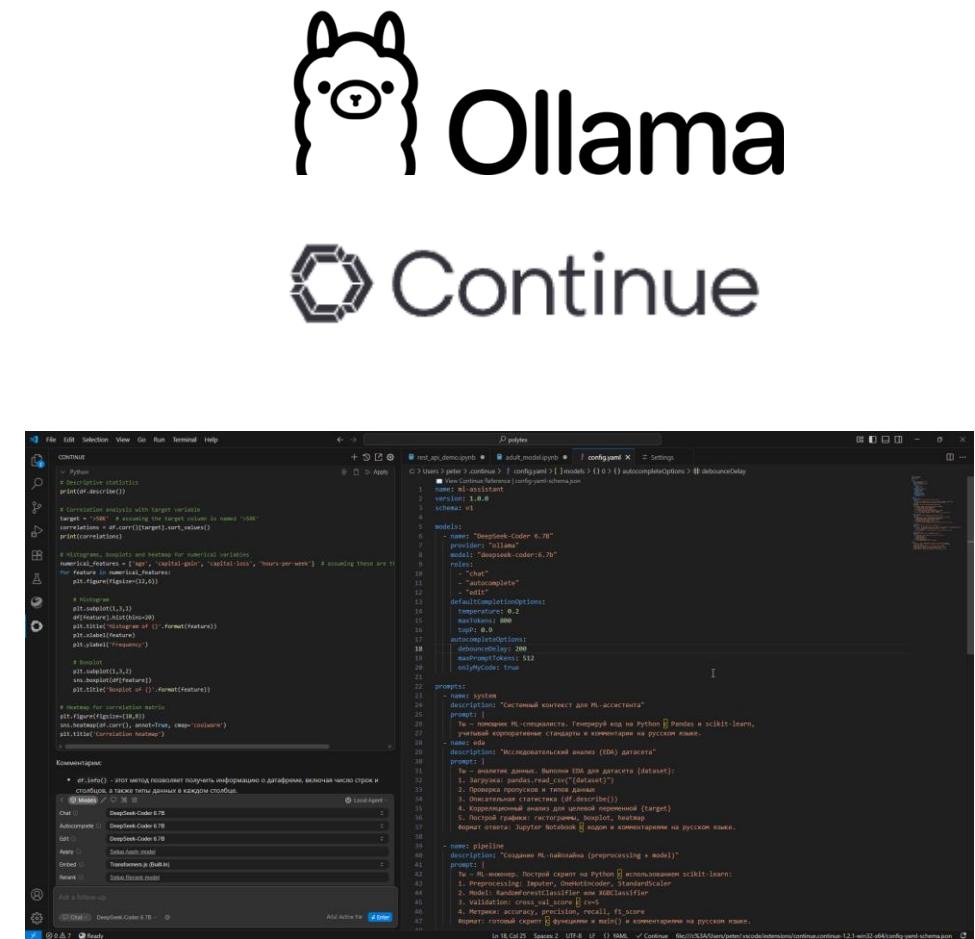
- есть бесплатный режим;
- экономия времени и сил на отладке;
- помощь в создании более эффективных ML-моделей;
- поддержка отечественных технологий, бесплатно, не требует VPN.



The screenshot shows the GigaCode interface running in a Microsoft Visual Studio Code window. The main area displays a Python script named `anomaly_detection.ipynb`. The code uses the `Faker` library to generate random product names and SKUs, then filters and scales the data, and finally uses a `LocalOutlierFactor` model to detect anomalies. A tooltip provides an explanation for the `generate_product` function. Below the code editor is a terminal window showing the output of the script's execution. A status bar at the bottom indicates the current file is `anomaly_detection.ipynb`, there are 0 errors and 0 warnings, and 1 file and 0 cells to analyze.

AI-ассистент на плагине Continue

- Open-source AI-ассистент для VS Code и JetBrains
- Поддержка множественных провайдеров LLM
- Гибкая конфигурация и настройка поведения
- Активное сообщество разработчиков



Фреймворк LightAutoML от Sber Tech

- 1. Один из самых практических AutoML (мнение экспертов Alfa Digital).**
- 2. Отечественная разработка, активное сообщество пользователей.**

Плюсы:

- Высокое качество моделей
- Гибкая конфигурация обучения
- Автоматический отчет
- Полноценная интерпретация (SHAP)
- Быстрый и долгий режимы

Минусы:

- Медленный отбор признаков
- Неполная документация (имеется русскоязычный тг-чат)



Первый релиз – май 2021.

Методическое сопровождение

Студенты переносили в Python
сделанный ранее кейс в
Loginom

- Примеры промптов – готовые формулировки
- Ожидаемые результаты – описание точного результата на каждом шаге
- Рефлексия по результатам работы

Шаг 4. Объединение профиля и события в выборку

Задача: соединить профили клиентов с разметкой события в единую выборку.

ПРИМЕР ПРОМПТА

Напиши функцию `create_training_sample(profile_df, events_df)`.

Функция должна:

1. Выполнить `inner join` профиля и событий по `'client'` (оставляем только клиентов, которые есть в обоих датафреймах)
2. Проверить, что целевая переменная `'event'` не содержит пропусков
3. Удалить служебные колонки с датами (они не нужны для моделирования)
4. Убедиться что в выборке есть оба класса для события (`True` и `False`)
5. Вернуть готовую выборку

На выходе: `DataFrame` где каждая строка = (профиль клиента + событие).

Ожидаемый результат

- Размер обучающей выборки: 39906 записей
- Распределение классов:
 - `False`: 32375 (81.2%)
 - `True`: 7531 (18.8%)
- Признаки: `['client', 'Recency', 'Frequency', 'Monetary', 'total_quantity', 'avg_check', 'total_unique_items', 'avg_items_per_visit', 'weekend_visits', 'amount_last_visit', 'event']`



Не забудьте добавить ваши 2 дополнительных признака

Промежуточные результаты

Критерии успешности проекта

90%

Доля магистрантов, успешно выполнивших все практические задания курса

90%

Доля магистрантов, планирующих использовать AutoML и генеративный ИИ в своих ВКР

100+

Количество слушателей онлайн-курса «Аналитика самообслуживания в Loginom» на открытом ресурсе

5+

Преподавателей СПбПУ планируют внедрить результаты проекта в свой учебный процесс

Промежуточный результат

- 88% студентов успешно выполнили все 7 практических работ I семестра
Из них 1 магистрант испытывал трудности с low-code парадигмой (опытный python-разработчик)
- 12% (2 чел.) отстают и имеют долги

В процессе

Анализ обратной связи

Что сработало

- Быстрое преодоление начальных барьеров входа в Python-программирование
- Мгновенная помощь при ошибках вместо поиска в документации
- Рекомендации по оптимизации кода и best practices
- Развитие навыка правильной формулировки задач (prompt engineering)

Ключевая компетенция

Студенты научились использовать AI-ассистента как инструмент усиления своих способностей, а не как замену обучению. Они корректно применяют code review, задают уточняющие вопросы, понимают логику сгенерированного кода.

Метрика успеха

Студенты выполнили практическую работу, повторив на Python весь сценарий, который ранее они делали в Loginom, в том числе большой блок по ETL — задачу, которая была бы недоступна большинству без помощи AI-ассистента.

Открытая часть проекта

Цикл видеолекций по эволюции аналитики данных и введению в машинное обучение

Следите за анонсами в телеграм-
канале вузов-участников
академической программы

Вводный курс по self-
service аналитике данных
на открытом ресурсе



Выводы

1. Главное – компетенция решить задачу, а не уметь писать код: в ML студент должен думать над данными, а не над синтаксисом; код – это инструмент, не профессия
2. Гибридный подход работает: платформа низкого кода + традиционный кодинг + AI ускоряют обучение аналитике данных
3. Выравнивание: Loginom нивелирует различия в подготовке студентов, позволяя всем сосредоточиться на логике, а не на языке программирования
4. AI-инструменты – помощники, не замена: LLM ускоряет разработку, но логика и решение задачи остаются в руках студента