

Требования: презентация на 5-7 минут(номера слайдов, шрифт!), в докладе должен быть один конкретный график или маленький пример кода

1	Голубева Анастасия	MLOps Engineer vs Data Scientist vs Analyst: зоны ответственности на проекте (Почему DS не должен деплоить модели в прод сам).
2	Григорян Луиза	Data Drift и Concept Drift: как убивают модели в реальной жизни (Кейсы: от кредитного скоринга до прогноза погоды).
3	Дущенков Даниил	Бритва Оккама в DS: почему простая линейная регрессия часто побеждает нейросеть (Эффективность, интерпретируемость, cost-benefit анализ).
4	Иванов Данил	Этика алгоритмов: дискриминация и «нечестные» признаки (Пример: гендерный дисбаланс в найме через ИИ или скоринг в США).
5	Лазутина Елена	Интерпретируемость «черных ящиков»: LIME и SHAP на пальцах (Как объяснить бизнесу, почему нейросеть отказала в кредите).
6	Лобаченко София	Проклятие размерности для кластеризации: почему евклидово расстояние перестает работать (Интуитивное объяснение для 100+ признаков).
7	Овсепян Нарине	DBSCAN против K-Means: когда форма кластера имеет значение (Кластеры-«сосиски» против кластеров-«шариков»).
8	Сиряева Галина	Оценка качества кластеризации без «правильного ответа»: Silhouette и индексы (Как понять, хорошо ли мы сгруппировали данные).
9	Терешонкова Мария	Алгоритмы поиска аномалий (Isolation Forest, LOF) (Как найти выброс без знания распределения).
10	Швец Наталья	Оценка качества кластеризации без «правильного ответа»: Silhouette и индексы (Как понять, хорошо ли мы сгруппировали данные).
11	Ищенко Полина	Работа с временными рядами: почему кросс-валидация "вслепую" — это читерство (Нельзя перемешивать прошлое и будущее).
12	Корж Андрей	Tiny Data: что делать, если у вас 100 наблюдений, а не 100 млн (Байесовские методы, синтез данных, очень простая модель).
13	Медоян Ирина	P-значения в мире ML: почему мы почти не используем t-тесты в продакшне (Различие между статистической значимостью и практической применимостью).
14	Мордюкова Алина	Автоматический поиск гиперпараметров (GridSearch, RandomSearch, Optuna) (Почему перебор сеткой — это зло, а байесовская оптимизация — добро).

15	Мышкова Ангелина	Воспроизводимость эксперимента (Reproducibility Crisis): Docker, DVC и фиксация random_seed (Как сделать так, чтобы модель работала на машине у коллеги).
16	Нарыжная Вероника	Parquet против CSV: почему формат хранения решает F1-скорость (Колоночное хранение и статистика по страницам).
17	Непошивайленко Мария	Dask и Modin: когда Pandas перестает влезать в RAM (Out-of-core вычисления и ленивые вычисления).
18	Полупанова Ирина	SQL в DS: не только SELECT, но и оконные функции для фич (Как сгенерировать скользящее окно в чистом SQL).
19	Тихоненко Мария	Feature Store: зачем переиспользовать признаки между моделями (Чтобы инженер и аналитик не считали одно и то же по-разному).
20	Фидрик Юлия	Проклятие размерности для кластеризации: почему евклидово расстояние перестает работать (Интуитивное объяснение для 100+ признаков).
21	Цыгикало Семен	DBSCAN против K-Means: когда форма кластера имеет значение (Кластеры-«сосиски» против кластеров-«шариков»).