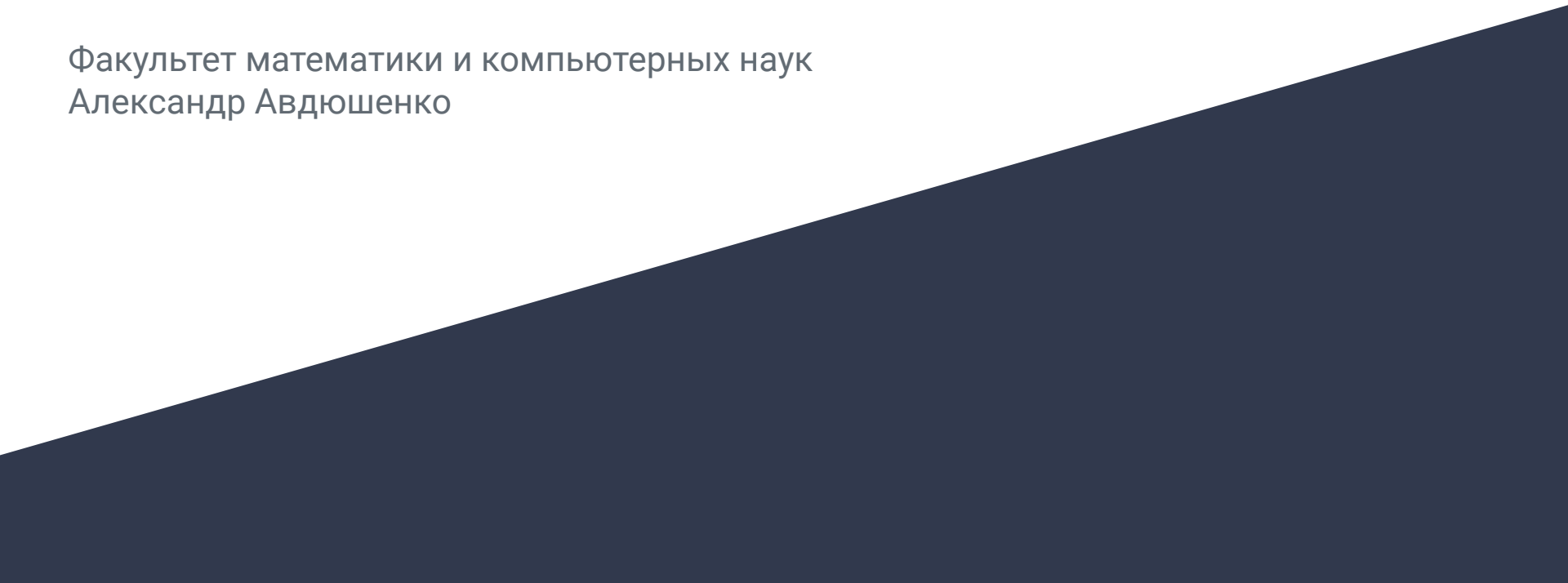


# ML тренировки в СПбГУ

Факультет математики и компьютерных наук  
Александр Авдюшенко

A dark blue diagonal gradient bar that starts from the bottom left and extends towards the top right, covering the lower half of the slide.

# Знакомство

- Записаться в листик
- У кого есть «золото»?
- У кого «серебро»?
- Любая медаль или сабмит?

# TODO

- [kaggle.com](https://kaggle.com) — регистрируемся
- [ods.ai](https://ods.ai) — присоединяемся (там небольшая форма)
- выбираем 1-2 соревнования для участия (см. план соревнований)
- сначала сабмитим индивидуально, потом объединяемся в команды
- в течение следующей недели — настраиваем выч. мощности от Botan Investments

# О соревнованиях

- Собственная (кросс)-валидация **очень** важна
- Видеть не только данные, но и процессы и задачу за ними
- Знайте свои инструменты (LightGBM, PyTorch, ...)
- Всегда после (или даже до) соревнования изучайте топовые решения
- Сложно – как fulltime работа
- Захватывает и в команде намного веселее

# О соревнованиях

- Не тратьте слишком много времени на гиперпараметры — лучше пробуйте разные идеи
- **Очень** полезно читать Discussion на Kaggle и иногда ODS
- Kernels (Notebooks) and <https://colab.research.google.com>
- Сразу пишите чистый воспроизводимый код (пайплайн)
- <http://mltrainings.ru/>
- [Бонус](#) новым мастерам

# Ссылки

- Супер полезный курс от top-kagglers  
<https://www.coursera.org/learn/competitive-data-science>
- Как правильно «фармить» Kaggle  
<https://habr.com/ru/company/ods/blog/426227/>
- Канал московских тренировок  
<https://www.youtube.com/channel/UCeq6ZIlvC9SVsfhfKnSvM9w/about>
- Images  
<https://github.com/lightforever/mlcomp> and  
<https://github.com/albu/albumentations>
- Tables <https://github.com/konodyuk/kts>