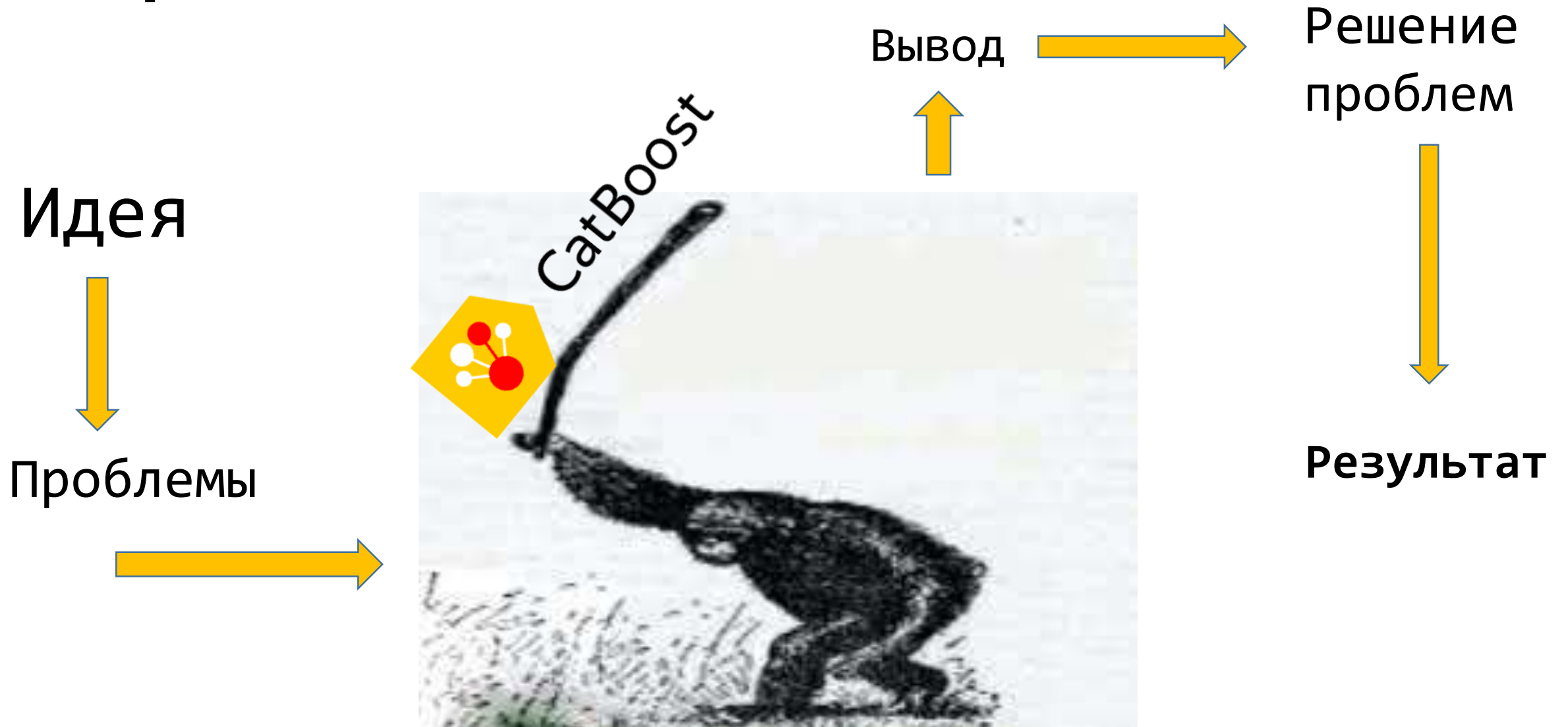


# Классификация сессий протоколов

Градиентный бустинг  
анализ признаков  
проблемы и их решения

**Команда:** Bebetto

# Pipeline



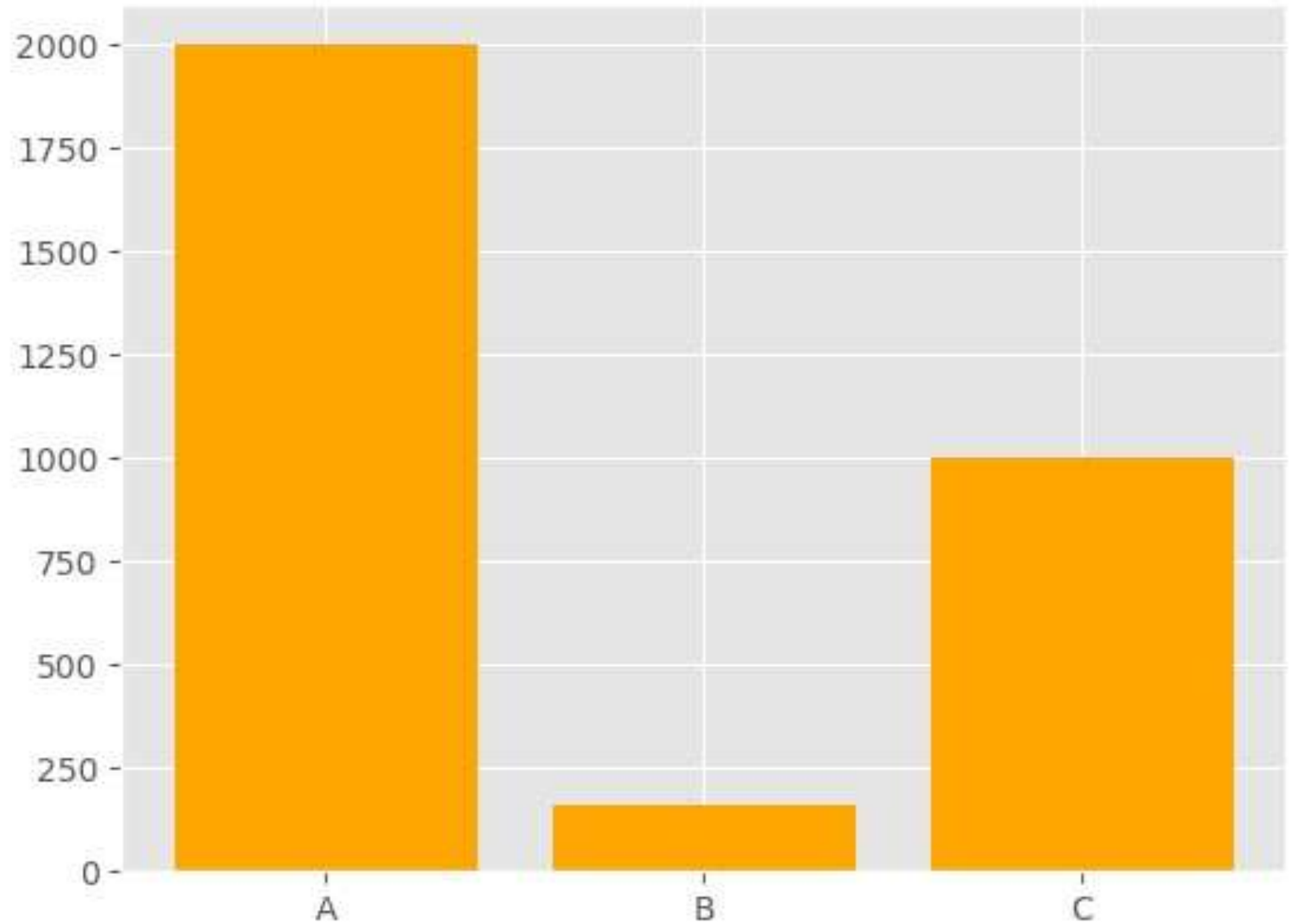
# Признаки и работа с ними

- Удаление ненужной информации – как не потерять в точности?
- Последовательность байт (Payload) – как извлечь полезную информацию?
- Несбалансированность классов – влияет ли на качество модели?
- Выделение ключевых признаков

# Количество сессий протоколов

Решение:

Балансировка  
датасета



# Ключевые признаки

