



Уральский
федеральный
университет

имени первого Президента
России Б.Н. Ельцина

Институт
фундаментального
образования

Автоматизация анализа текстов партий

Докладчики

Быков Иван(ФО-360001)

Куратор

Шibaев Вячеслав Алексеевич

Введение



MAX-PLANCK-INSTITUT
FÜR AUSLÄNDISCHES ÖFFENTLICHES RECHT
UND VÖLKERRECHT

Цели

- В достаточной мере освоить язык программирования C# и Accord framework;
- Используя различные методы подготовки и обработки данных, добиться точности SVM в 42%;
- Попытаться превзойти этот процент точности с помощью новых методов.

Используемые технологии



C# 7.0



Microsoft visual
studio 2017



Accord Framework
v3.8

- Изучение теории
- Подготовка текста
- Получение tf/idf матрицы
- Обучение классификатора
- Проверка точности на разных выборках

Получена необходимая точность

Векторизация	Размер выборки	Удаление стоп-слов и цифр	Точность, %
Да	Большая	Нет	20,36
Да	Большая	Да	44,77
Да	Малая	Да	45,83



Уральский
федеральный
университет

имени первого Президента
России Б.Н.Ельцина

Институт
фундаментального
образования

Спасибо за внимание!