Здравствуйте, мы представляем команду CyberBorsch AA. Ее представители Данил Авагян, Сахинов Сергей, Евгений Егоров, Сергей Никишов.

Перед нами стояла задача сделать быструю и удобную программу для сортировки файлов. Мы распределили задачи и начали работать. Сергей Сахинов и Сергей Никишов отвечали за дизайн, документацию, а также помогали в начальной разработке классификатора. Евгений Егоров и Данил Авагян отвечали за код. После распределения обязанностей, мы рассмотрели несколько возможных вариантов работы программы и нашли подходящий.

Для классификации мы представляем слова в виде вектора при помощи метри TF-EDF и вычисляем косинусное расстояние между вектором текста и вектором темы. Чем ближе получившееся число к единице - тем больше вероятность того, что файлы относятся к одной теме.

Для оптимизации кода мы решили пойти не стандартным путем, а именно использовать сразу два языка программирования: python и c++. Это, в свою очередь, помогло нам ускорить обработку файлов, что в конечном итоге ускорило работу самой программы.

1. Представление участников
2. Что за проект, какие цели были поставлены
3. метод классификации
4. соединение python и c++, почему
5. Сказ про код великий
6. …
7. Рассказ про функционал, при нажатии на HELP рассказать про документацию
8. Показать как работает на файлах