



«Слепая» стилометрия

почему Delta не отличает
Мандельштама от
нейромандельштама

Борис Орехов (ИРЛИ РАН)

ЭЛЕКТРОННАЯ ЛИРА

Как и зачем
искусственный
интеллект пишет стихи?

Издательство

Европейского университета

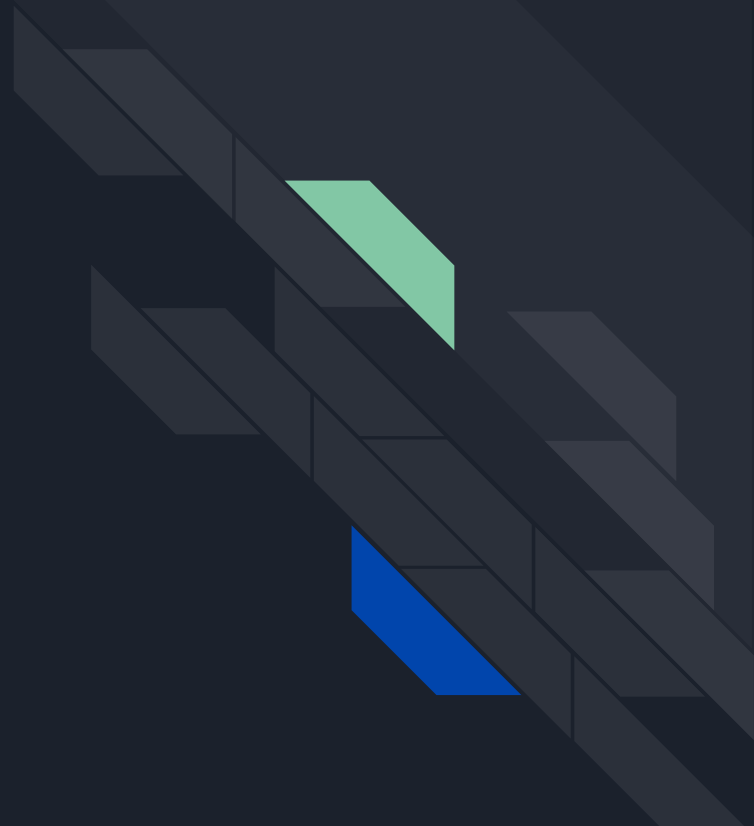
в Санкт-Петербурге

ISBN 978-5-94380-407-6

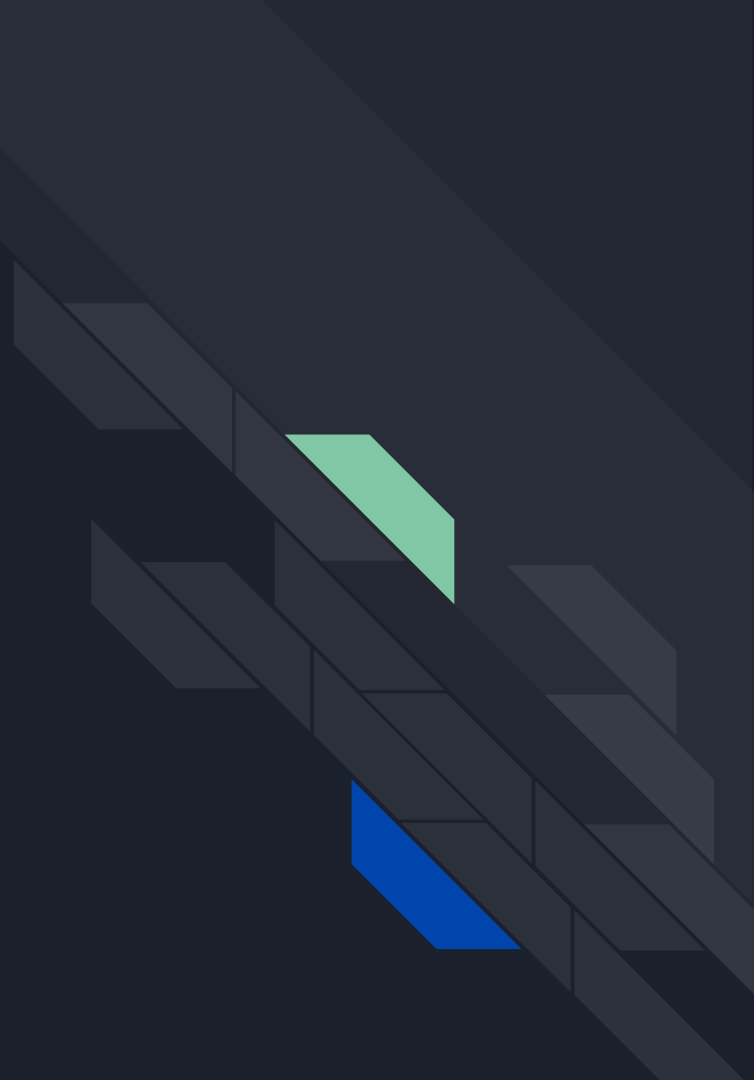
Год издания 2026



Главное, что интересует в
поэтической продукции
нейросетей, это их
сходство с
оригинальными текстами



Эксперимент со стилем





Проблема ручной и машинной оценки

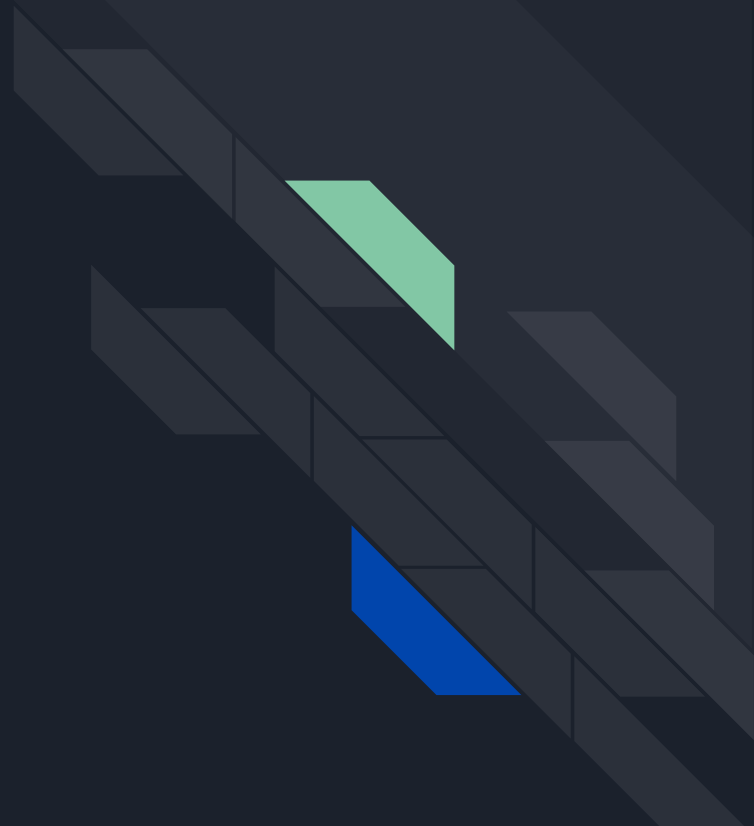
- Ручная оценка — не единственный путь.
- Существуют методы, где компьютер сам определяет автора.
- Среди них есть достаточно надёжные.
- Но смогут ли они показать, насколько хорошо нейросеть воспроизводит стиль?



Почему эксперимент может быть «игрой в одни ворота»

- Нейросети обучаются на текстах через сложную статистику.
- Методы определения авторства тоже основаны на статистике.
- Если обучать статистическую модель, а потом статистически её проверять, она будет «смотреться в зеркало».

Речь о стихах, которые
писали нейросети
предыдущего
современному
поколения: LSTM





Примеры

Температура 0,5

И за ней веселый поверник,
И напеленной какой-то городах
Неусухие деле разной,
Заметил на луне прозрачное вода,
Или без вечного переплатанный воздух
проницами,
И в корибере достоин коленях.
Он вода не дышать воздух проницами
положился
В передавали в нем на свете нашей
приполоматый

Температура 1

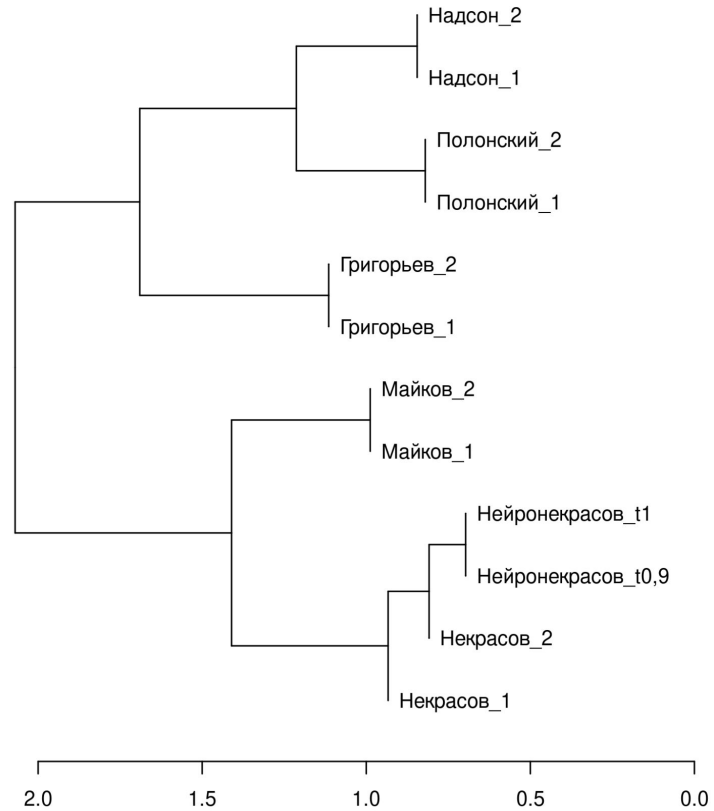
Я марю двери в смолой ногам,
На ямре других на пряда стакой
Висит вороха из из твой жизни мучил,
Оборонулся поднимал мил. В
мостренский увидет золотий!
Так в стеклян иой Эльгательноты.
И для двух должна, как бывуйство.
Я отобразам игла огнем.
И же так как хилой розовых плечах.



Как работает метод Delta

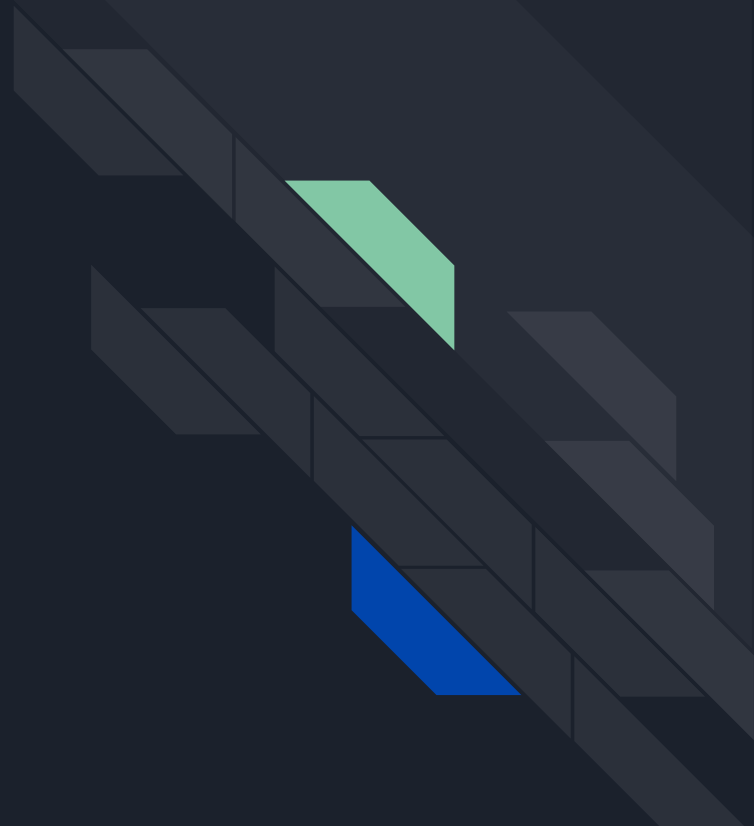
- Предложен Джоном Берроузом в 2002 году.
- Самый популярный способ количественной атрибуции в digital humanities.
- Шаги:
 - рассчитать частотности всех слов в коллекции;
 - выбрать 100–300 самых частотных слов;
 - сравнивать тексты попарно по этим словам;
 - усреднять различия и получать меру близости.

Некрасов Cluster Analysis

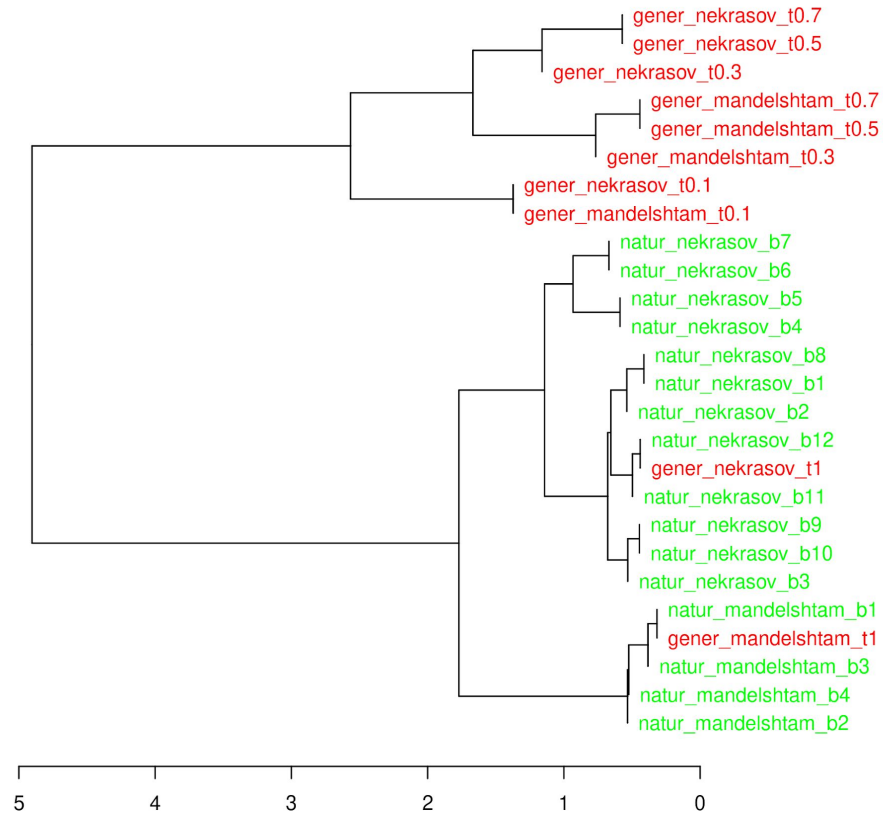


100 MFW Culled @ 0%
Classic Delta distance

Температура — параметр генерации: чем выше, тем свободнее и разнообразнее текст, чем ниже — тем предсказуемее и однообразнее.



char_1stm Cluster Analysis



100 MFW Culled @ 0%
Classic Delta distance



Парадокс и вывод

- Нейросеть выучивает статистику оригинала и обманывает методы атрибуции.
- Delta учитывает только 100 частотных слов — редкие «новые» слова не мешают.
- Для человека текст бессвязен, для статистики — «больше Мандельштам, чем сам Мандельштам».
- С оригиналом Delta связывает только нейромандельштама с высокой температурой.
- При 0,7 и меньше метод различает генерацию и оригинал.
- Человек смотрит на связность и логику, машина — на распределения.