**Лабораторная работа №3.**

**RNN в обработке текстов.**

**Задание 1.** Решите задачу расшифровки сообщения с помощью RNN. Представьте, что вам даны сообщения, зашифрованные с помощью шифра Цезаря, являющимся одним из самый простых шифров, в криптографии.

Шифр цезаря работает следующим образом: каждая буква исходного алфавита сдвигается на K символов вправо:

Пусть нам дано сообщение: message="RNN IS NOT AI", тогда наше шифрование выполняющиеся по правилу f, с K=2, даст нам результат: f(message, K) = TPPAKUAPQVACK.

Для удобства можно взять символы только одного регистра в нашей имплементации, и сказать, что все буквы не английского алфавита будут отмечены как прочерк "-"

**Задание 2.** Примените RNN к классификации текстов из лабораторной работы №1 (на несбалансированных и балансированных классах). Сравните результаты с классическими методами машинного обучения. Сделайте выводы!

**Задание 3.** Для настоящего корпуса собраний сочинений Федора Михайловича Достоевского сделайте следующие задания:

Сколько слов и предложений в датасете?

Каковы частоты для слов «бесы», «семья», «брат». Введите значение, округленное до 5 знаков?

Сгенерируйте текст, который состоит из слов с вероятностями появления от 0. 85 до 0.95 и итоговое предложение состоит из 10 слов.