Учреждение образования

«Белорусский государственный технологический университет»

**Лабораторная работа №2**

Элементы теории информации. Информативность данных в различных кодировках

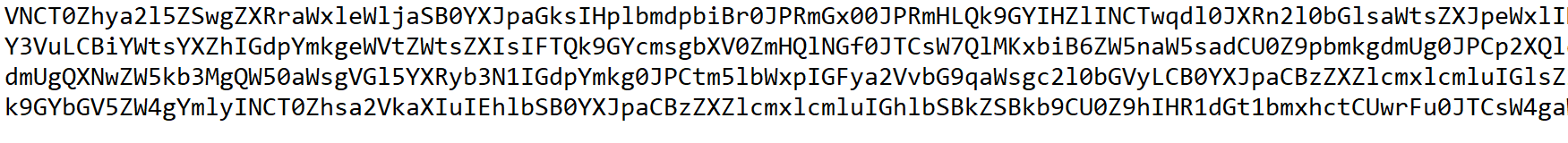
Выполнил:

Студент 3 курса 5 группы ФИТ

Коршун Никита Игоревич

2024

**Задание 1.** Для конвертации в *base64* было создано приложение. В качестве входного документа использовался текстовый файл с предыдущей лабораторной работы с текстом на турецком языке (латиница).

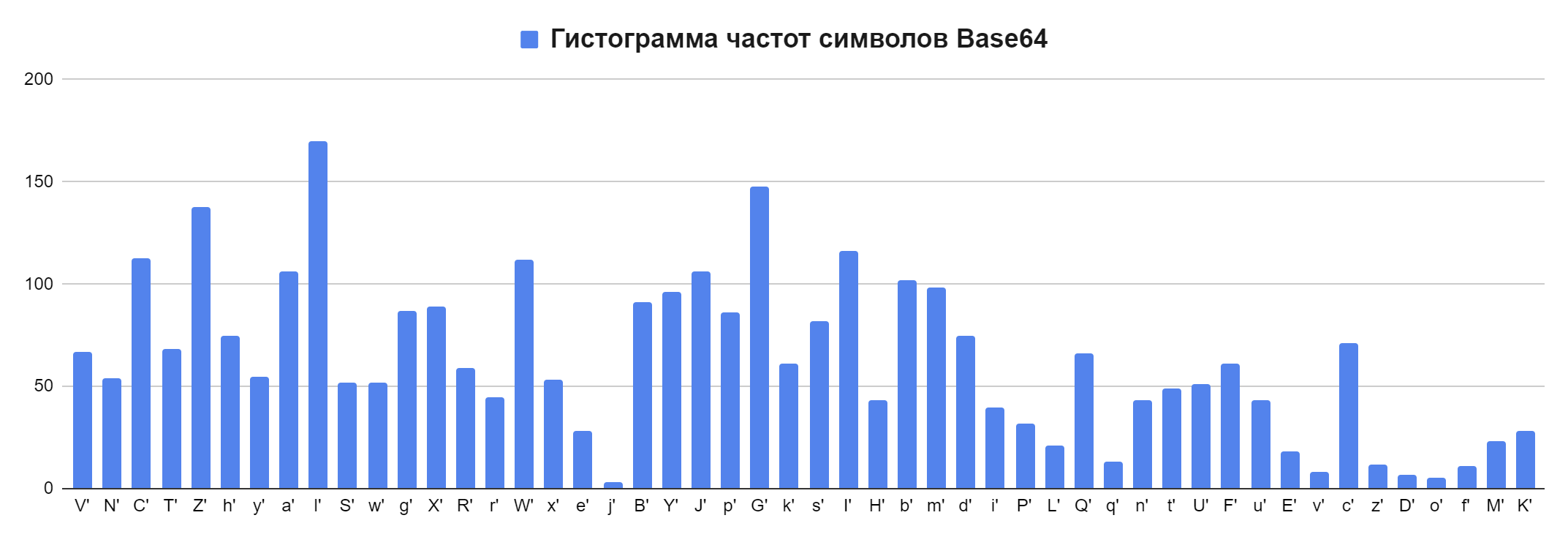


Листинг 1.1 – Результат преобразования текста в base64

**Задание 2.** Для реализации за основу были взяты функции из предыдущей лабораторной работы, после небольшой доработки, были добавлены возможность вычисления энтропии Хартли и избыточности алфавита.



Листинг 2.1 – Гистограмма частот символов турецкого языка



Листинг 2.2 – Гистограмма частот символов *Base64*

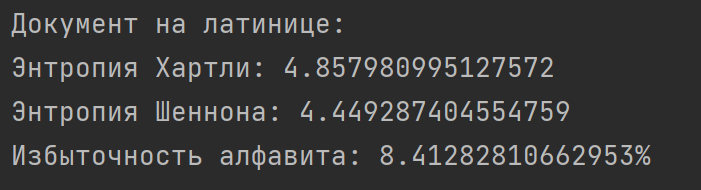


Рисунок 2.1 – Результат вычислений для латинского алфавита

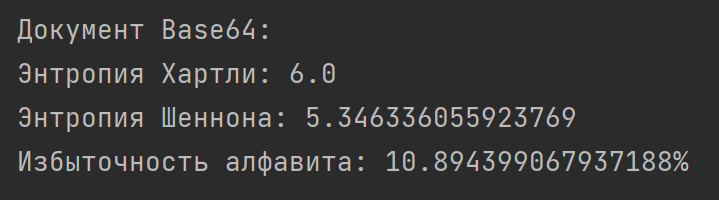
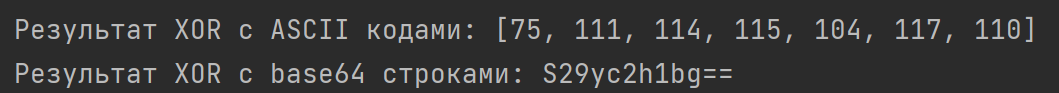


Рисунок 2.2 – Результат вычислений для base64

**Задание 3.** Для реализации была создана функция на языке Python, которая позволяет выполнять операции XOR. Если имеется строка с меньшей длиной, она дополняется нулями.



Листинг 3.1 – Результат вычислений XOR

**Вывод:** В данном случае, избыточность турецкого сообщения составляет 3.9%. Это означает, что алфавит турецкого сообщения использует примерно на 3.9% больше бит, чем необходимо для передачи информации без потерь.

В случае с *Base64* избыточность алфавита составляет 5,1%. Это означает, что алфавит *Base64* использует примерно на 5,1% больше бит, чем необходимо для передачи информации без потерь. Исходя из этого, использование латинского языка для передачи и хранения более эффективно, чем использование *Base64*.