МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ

Учреждение образования «БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ

ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ»

**ОТЧЕТ ПО ЛАБОРАТОРНОЙ РАБОТЕ №2 НА ТЕМУ:**

**Элементы теории информации. Параметры и характеристики дискретных информационных систем**

Выполнила студентка 3 курса 4 группы

Сятковская Екатерина Дмитриевна

Минск 2023

**Задание 1:** Рассчитать энтропию двух алфавитов: один на основе латиницы, другой – на кириллице

Энтропию алфавита по К. Шеннону рассчитывают по следующей формуле:

Для выполнения задания были выбраны белорусская кириллица и белорусская латиница. Для расчёта использовался следующий код на языке Python:

def text\_reader(path):  
 with open(path**,** 'r') as f:  
 text = f.read().lower()  
 return text  
  
def letters\_dict(text):  
 letters\_dict = collections.defaultdict(int)  
 for char in text:  
 if char.isalpha() or char.isdigit():  
 letters\_dict[char] += **1** return letters\_dict  
  
def probs(text):  
 letter\_dict = letters\_dict(text)  
 letters\_probs = {}  
 for char**,** count in letter\_dict.items():  
 letters\_probs[char] = count / sum(letter\_dict.values())  
 return letters\_probs  
  
def entropy(text):  
 entropy = **0** letters\_probs = probs(text)  
 for prob in letters\_probs.values():  
 entropy += prob \* math.log(prob**, 2**)  
 return -entropy

Листинг 1 – задание 1

В результате выполнения в консоль выведутся значения энтропии белорусской кириллицы и белорусской латиницы.

Гистограмма частоты появления символов в тексте кириллицы:

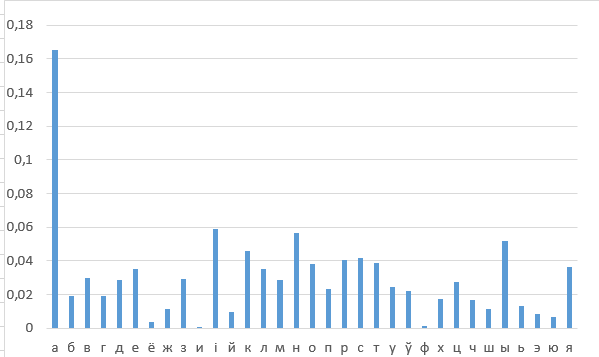


Рис 1 – Гистограмма текста кириллицы

Гистограмма частоты появления символов в тексте латиницы:

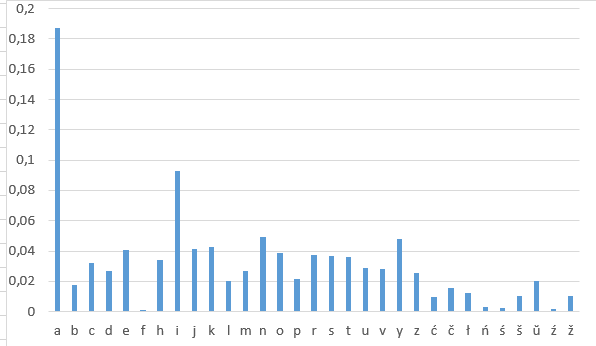


Рис 2 – Гистограмма текста латиницы

**Задание 2:** Рассчитать энтропию бинарного алфавита на основе входного файла

Дополним разработанный в предыдущем задании код следующий метод:

def convert\_to\_ascii(text):  
 ascii\_text = ''  
 for char in text:  
 if char.isalpha():  
 ascii\_text += bin(ord(char))[**2**:]  
 return ascii\_text

Листинг 2 – задание 2

Данный метод позволяет перевести исходные входные файлы в бинарный алфавит, расчет энтропии происходит с помощью методов, описанных в задании 1.

**Задание 3:** Используя значения энтропии алфавитов, полученные в заданиях 1 и 2, подсчитать количество информации в сообщении, состоящем из собственных имени, фамилии и отчества

Дополним разработанный в предыдущем задании код следующим методом:

def quantity\_of\_information(entropy**,** text):  
 return entropy \* len(text)

Листинг 3 – задание 3

**Задание 4:** выполнить задание 3 при условии, что вероятность ошибочной передачи единичного бита сообщения составляет: 0.1, 0.5, 1.

Дополним разработанный в предыдущем задании код следующим методом:

def mistake\_quantity(mistake\_prob**,** text**,** entropy):  
 return len(text) \* (entropy - (-mistake\_prob \* math.log(mistake\_prob**, 2**) - (**1** - mistake\_prob) \* math.log(**1** - mistake\_prob**, 2**)))

Листинг 4 – задание 4

Результат работы итогового кода:

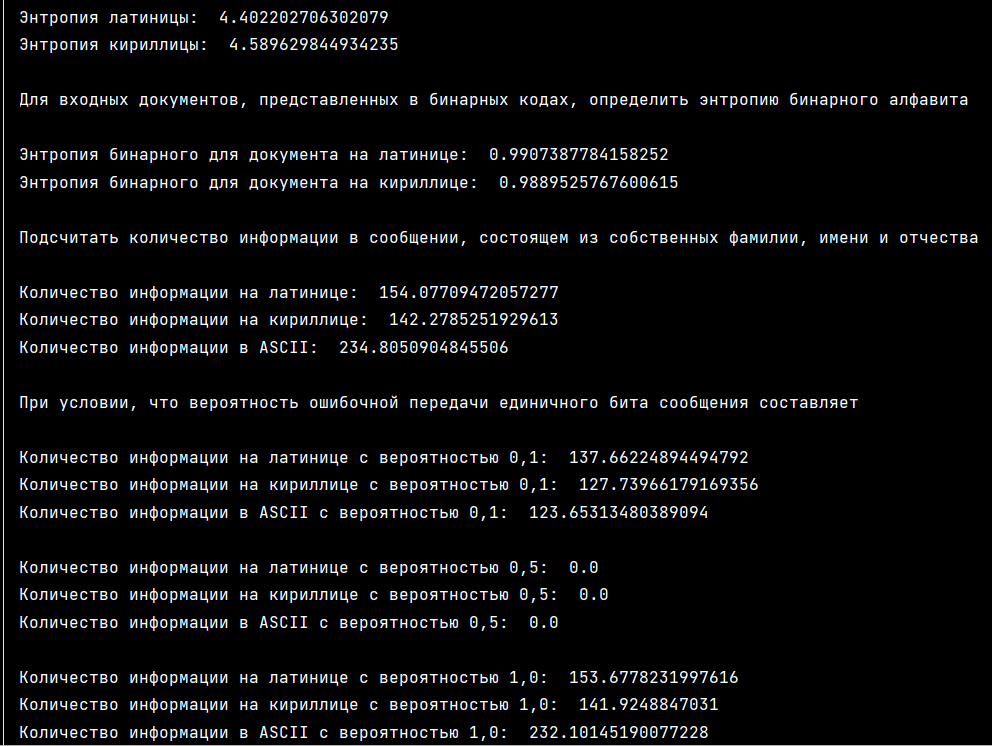


Рис 3 - Результат работы программы

**Вывод:** в ходе лабораторной работы были изучены методики расчёта энтропии алфавитов и количества информации в сообщениях на основе этих алфавитов. Приобретённые знания были применены для латышского и болгарского алфавитов.