МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное

учреждение высшего образования

«ЧЕРЕПОВЕЦКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

|  |
| --- |
| Институт информационных технологий |
| Кафедра математического и программного обеспечения |
| Теория информации |
|  |

ЛАБОРАТОРНАЯ РАБОТА №1

«Количественная оценка информации»

|  |  |
| --- | --- |
| Исполнитель: | студент  группы 1ПИб-02-3оп-23  Богданов  Ренат Алексеевич |
| Руководитель: | Ганичева Оксана Георгиевна |
| Оценка: |  |
| Подпись: |  |

2024, Череповец

Задание

Написать текст, содержащий краткую информацию об одном из основоположников теории информации К. Шенноне. (Рекомендуется взять один из «10 уроков из жизни основоположника теории информации») Текст должен содержать не менее 7-10 строк, отформатированных по ширине страницы.

Часть 1: использование алфавитного (структурного) подхода

1. Подсчитать количество информации, содержащейся в заданном тексте.
2. Решение задачи оформить в таблице в программе MS Excel. Считать, что в алфавит символов входят все буквы русского языка.
3. Отчет по работе оформить в программе MS Word, с использованием вставки с решением из программы MS Excel (создать комплексный документ).

Часть 2: использование содержательного(вероятностного) подхода

1. Посчитать количество информации, содержащейся в заданном тексте, используя вероятностный подход.
2. Решение задачи оформить в программе MS Excel.
3. Отчет по работе оформить в программе MS Word, с использованием вставки с решением из программы MS Excel (создать комплексный документ).

Ход работы

Текст:

“Является основателем теории информации, нашедшей применение в современных высокотехнологических системах связи. Предоставил фундаментальные понятия, идеи и их математические формулировки, которые в настоящее время формируют основу для современных коммуникационных технологий. В 1948 году предложил использовать слово «бит» для обозначения наименьшей единицы информации (в статье «Математическая теория связи»). Кроме того, понятие энтропии было важной особенностью теории Шеннона. Он продемонстрировал, что введённая им энтропия эквивалентна мере неопределённости информации в передаваемом сообщении. Статьи Шеннона «Математическая теория».”

I. Использование алфавитного (структурного) подхода

1. Находим количество символов [k] в тексте. Текст состоит из 640 символов.
2. Найти мощность алфавита [N]. Количество уникальных символов равно 61.
3. По формуле Хартли (1.1) рассчитать количество информации на один символ.

|  |  |  |
| --- | --- | --- |
|  |  | (1.1) |

Количество информации на один символ:**5,930737**

1. По формуле (1.2) находим количество информации, приходящееся на все символы.

|  |  |  |
| --- | --- | --- |
|  |  | (1.2) |

Количество информации: **3795,672**

Решение, выполненное с помощью Excel, приведено в таблице 1.

Таблица 1

Структурный подход

|  |  |  |  |
| --- | --- | --- | --- |
| k= | 640 | *‘=ДЛСТР(A1*) |  |
| N= | 61 | *‘=LOG(B4; 2)* |  |
| i= | 5,930737 |  | i=log2 N |
| **I=** | **3795,672** | *‘=B3\*B5* | I=K log2 N |

II. Использование содержательного (вероятностного) подхода

1. Находим вероятность встречи для каждого символа по формуле (2.1).

|  |  |  |
| --- | --- | --- |
|  | , | (2.1) |

где n – количество определённого символа в тексе, N – общее количество символов в тексте.

1. По формуле Шеннона (2.2) находим количество информации.

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  | (2.2) |

Таблица 2

Вероятностный подход

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| P | n(i) - стр. | p(i) | log2 | p(i)\*log2 |
| а | 33 | 0,0515625 | -4,27753 | -0,220560346 |
| б | 5 | 0,0078125 | -7 | -0,0546875 |
| в | 25 | 0,0390625 | -4,67807 | -0,182737184 |
| г | 4 | 0,00625 | -7,32193 | -0,045762051 |
| д | 13 | 0,0203125 | -5,62149 | -0,114186483 |
| е | 58 | 0,090625 | -3,46395 | -0,313920206 |
| ё | 2 | 0,003125 | -8,32193 | -0,026006025 |
| ж | 2 | 0,003125 | -8,32193 | -0,026006025 |
| з | 4 | 0,00625 | -7,32193 | -0,045762051 |
| и | 56 | 0,0875 | -3,51457 | -0,307525153 |
| й | 4 | 0,00625 | -7,32193 | -0,045762051 |
| к | 11 | 0,0171875 | -5,8625 | -0,100761658 |
| л | 17 | 0,0265625 | -5,23447 | -0,139040483 |
| м | 27 | 0,0421875 | -4,56704 | -0,192672025 |

Продолжение таблицы 2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| P | n(i) - стр. | p(i) | log2 | p(i)\*log2 |
| н | 49 | 0,0765625 | -3,70722 | -0,283833897 |
| о | 60 | 0,09375 | -3,41504 | -0,320159766 |
| п | 11 | 0,0171875 | -5,8625 | -0,100761658 |
| р | 27 | 0,0421875 | -4,56704 | -0,192672025 |
| с | 25 | 0,0390625 | -4,67807 | -0,182737184 |
| т | 36 | 0,05625 | -4,152 | -0,233550174 |
| у | 6 | 0,009375 | -6,73697 | -0,063159052 |
| ф | 6 | 0,009375 | -6,73697 | -0,063159052 |
| х | 8 | 0,0125 | -6,32193 | -0,079024101 |
| ц | 5 | 0,0078125 | -7 | -0,0546875 |
| ч | 6 | 0,009375 | -6,73697 | -0,063159052 |
| ш | 5 | 0,0078125 | -7 | -0,0546875 |
| щ | 2 | 0,003125 | -8,32193 | -0,026006025 |
| ъ | 0 | 0 | 0 | 0 |
| ы | 8 | 0,0125 | -6,32193 | -0,079024101 |
| ь | 7 | 0,0109375 | -6,51457 | -0,071253144 |
| э | 3 | 0,0046875 | -7,73697 | -0,036267026 |
| ю | 2 | 0,003125 | -8,32193 | -0,026006025 |
| я | 19 | 0,0296875 | -5,074 | -0,150634392 |
| 1 | 1 | 0,0015625 | -9,32193 | -0,014565513 |
| 2 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 0 | 0 |
| 4 | 1 | 0,0015625 | -9,32193 | -0,014565513 |
| 5 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 |
| 8 | 1 | 0,0015625 | -9,32193 | -0,014565513 |
| 9 | 1 | 0,0015625 | -9,32193 | -0,014565513 |

Продолжение таблицы 2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| P | n(i) - стр. | p(i) | log2 | p(i)\*log2 |
| 0 | 0 | 0 | 0 | 0 |
| ! | 0 | 0 | 0 | 0 |
| " | 0 | 0 | 0 | 0 |
| № | 0 | 0 | 0 | 0 |
| ; | 0 | 0 | 0 | 0 |
| % | 0 | 0 | 0 | 0 |
| : | 0 | 0 | 0 | 0 |
| ? | 0 | 0 | 0 | 0 |
| \* | 0 | 0 | 0 | 0 |
| ( | 1 | 0,0015625 | -9,32193 | -0,014565513 |
| ) | 1 | 0,0015625 | -9,32193 | -0,014565513 |
|  | 71 | 0,1109375 | -3,17218 | -0,351913827 |
| + | 0 | 0 | 0 | 0 |
| - | 0 | 0 | 0 | 0 |
| . | 6 | 0,009375 | -6,73697 | -0,063159052 |
| / | 0 | 0 | 0 | 0 |
| = | 0 | 0 | 0 | 0 |
| « | 3 | 0,0046875 | -7,73697 | -0,036267026 |
| » | 3 | 0,0046875 | -7,73697 | -0,036267026 |
| , | 5 | 0,0078125 | -7 | -0,0546875 |
|  | **640** | **1** |  | **4,525858424** |
|  | *=СУММ(B12:B73)* | *=СУММ(C12:C73)* |  | *=-1\*СУММ(E12:E73)* |
| **I=** | **2896,549391** | =E74\*B74 |  |  |

Таблица 3

Итоги

|  |  |
| --- | --- |
| I по Хартли | **3795,671896** |
| I по Шеннону | **2896,549391** |

Вывод

В ходе выполнения работы были освоены структурный и вероятностный методы оценивания количества информации. При сравнение итоговых значений (см. табл. 3), можно сделать вывод о том, что значения не совпадают. Это объясняется тем, что структурный подход не учитывает вероятность исходов события, а вероятностный подход учитывает как количество возможных исходов, так и их вероятности.

В результате структурный подход дает верхнюю границу количественной меры информации, а вероятностный подход дает более точное измерение, учитывающее вероятность исходов.