

Ψηφιακές Επικοινωνίες 1

Όνοματεπώνυμο : Νικόλας Φιλιππάτος

ΑΜ: 1072754

Εργασία: 1η

Ημερομηνία: November 23, 2023

Ερωτημα 1

Χρησιμοποιείτε το παρακάτω ποίημα να υπολογίσετε και να απαντήσετε στα εξής:

```

Livin' easy
Lovin' free
Season ticket on a one way ride
Askin' nothin'
Leave me be
Takin' everythin' in my stride
Don't need reason
Don't need rhyme
Ain't nothin' that I'd rather do
Goin' down
Party time
My friends are gonna be there too
I'm on the highway to hell
On the highway to hell
Highway to hell
I'm on the highway to hell
No stop signs
Speed limit
Nobody's gonna slow me down
Like a wheel
Gonna spin it
Nobody's gonna mess me around
Hey satan
Payin' my dues
Playin' in a rockin' band
Hey mumma
Look at me
I'm on the way to the promised land
I'm on the highway to hell
Highway to hell
I'm on the highway to hell
Highway to hell
Don't stop me
I'm on the highway to hell
On the highway to hell
Highway to hell
I'm on the highway to hell(Highway to hell) I'm on the highway to hell
(Highway to hell) highway to hell
(Highway to hell) highway to hell
(Highway to hell)
And I'm goin' down
All the way
I'm on the highway to hell

```

1. Την αυτοπληροφορία των L, h, l, H, s, n και w .

1. Την αυτοπληροφορία των L, h, l, H, s, n και w .

Η αυτοπληροφορία των γραμμάτων βρίσκεται απο τον τυπο $I(x_i) = \log_2 \left(\frac{1}{P(x_i)} \right) = -\log_2 (P(x_i))$, οπου $P(x_i)$ είναι η πιθανότητα εμφανισης του γραμματος, και βρίσκεται απο τον τυπο $P(x) = \frac{\text{Number of Letter}}{\text{Number of all the characters}}$, το πηλικο του αριθμου των γραμμάτων σε ολο το τραγουδι N=930.

Για τον Υπολογισμό του παρακάτω πίνακα αξιοποιήθηκε το παρακάτω [python πρόγραμμα](#).

Letter	Count	P = Count/N	I= -log2(P)
L	5	0.005376344086021506	7.5391588111108031
h	73	0.07849462365591398	3.6712623471153765
l	47	0.05053763440860215	4.306498054317756
H	10	0.010752688172043012	6.5391588111108031
s	19	0.02043010752688172	5.613159392551808
n	56	0.060215053763440864	4.05373198393779
w	28	0.030107526881720432	5.05373198393779
Number of Letters	751		
Number of Symbols	179		
Number of Characters	930		

Κώδικας

Ο κωδικας αναπτυχθηκε με σκοπο τα αποτελεσματα να καταγραφονται σε αρχεια markdown (md) για να γίνει αυτοματα η μορφοποιηση της αναφορας

```
#!/bin/python

from collections import defaultdict
import math

def getting_info_from_file(file, individual=False):
    """This function will return a dictionary with the frequency of each letter in the file"""
    letter_frequency = defaultdict(int)
    with open("poem.txt", "r") as file:
        for line in file:
            for letter in line:
                letter_frequency[letter] += 1
    letter_frequency["newLine"] = letter_frequency.pop("\n")
    return letter_frequency

def letters_of_interest_frequency(letters_of_interest, letter_frequency):
    """This function will return a dictionary with the frequency of each letter of interest of a given dictionary"""
    return {key: letter_frequency[key] for key in letters_of_interest}

def number_of_symbols(letter_frequency):
    """Returns the number of symbols in the dictionary"""
    summ = [
        letter_frequency[key] for key in letter_frequency.keys() if not key.isalpha()
    ]
    return sum(summ)

def print_table(dict_of_interest, letter_frequency, individual=False):
```

```

# total number of letters in the file
N = sum(letter_frequency.values())
n_symbols = number_of_symbols(letter_frequency)
n_letters = N - n_symbols

# beautifying the Markdown table
ln = len("Number of Characters") + 2

# the table headers
result = [
    f'| {"Letter":^{\ln}} | {"Count":^{\ln}} | {"P = Count/N":^{\ln}} | {"I= -log2(P)":^{\ln}} | ',
    f'| {"-":^{\ln}} | {"-":^{\ln}} | {"-":^{\ln}} | {"-":^{\ln}} | ',
]

for key, value in dict_of_interest.items():
    information = f"{(-1*math.log2(value/N))}"
    probability = f"{value/N}"

    result.append(
        f"| {key:^{\ln}} | {value:^{\ln}} | {probability:^{\ln}} | {information:^{\ln}} | "
    )

result.append(
    f'| {"Number of Letters":^{\ln}} | {N:^{\ln}} | {"":^{\ln}} | {"":^{\ln}} | '
)
result.append(
    f'| {"Number of Symbols":^{\ln}} | {n_symbols:^{\ln}} | {"":^{\ln}} | {"":^{\ln}} | '
)
result.append(
    f'| {"Number of Characters":^{\ln}} | {n_letters:^{\ln}} | {"":^{\ln}} | {"":^{\ln}} | '
)
return result

def sharing_data(letters_of_interest, poem_file, individual=False):
    # dictionary with the frequency of each letter in the file
    letter_frequency = getting_info_from_file(poem_file, individual)

    # dictionary with the frequency of each letter of interest
    dict_of_interest = letters_of_interest_frequency(
        letters_of_interest, letter_frequency
    )

    return letters_of_interest, letter_frequency, dict_of_interest

def save_to_file(file_name, result):
    with open(file_name, "w") as file:
        file.write("\n\n")
        file.write("\n".join(result))

def main(individual=False):
    letters_of_interest, letter_frequency, dict_of_interest = sharing_data(
        ["L", "h", "l", "H", "s", "n", "w"], "poem.txt", individual
    )

    # printing the table
    result = print_table(dict_of_interest, letter_frequency, individual)

    print("\n".join(result))

    save_to_file("assignment-1-1-code-result.md", result)

```

2. Την εντροπία της γλώσσας αυτού του κειμένου.

2. Την εντροπία της γλώσσας αυτού του κειμένου.

Η συνολική εντροπία της γλώσσας βρίσκεται απο τον τυπο: $H(X) = - \sum_{i=1}^N p_i \log_2(p_i)$

Θα βρούμε την πιθανοτητα εμφανισης και την αυτοπληροφορια για καθε χαρακτηρα του ποιηματος (εκτος του \n)

Letter	Count	P = Count/N	I= -log2(P)
L	5	0.005376344086021506	7.539158811108031
i	48	0.05161290322580645	4.2761244052742375
v	4	0.004301075268817204	7.861086905995394
n	56	0.060215053763440864	4.05373198393779
'	29	0.03118279569892473	5.003105910867822
	142	0.15268817204301074	2.711339786490712
e	74	0.07956989247311828	3.651633540366444
a	48	0.05161290322580645	4.2761244052742375
s	19	0.02043010752688172	5.613159392551808
y	36	0.03870967741935484	4.6911619045530815
o	65	0.06989247311827956	3.8387190929669393
f	2	0.002150537634408602	8.861086905995395
r	15	0.016129032258064516	5.954196310386875
S	2	0.002150537634408602	8.861086905995395
t	57	0.06129032258064516	4.028196891830652
c	2	0.002150537634408602	8.861086905995395
k	6	0.0064516129032258064	7.2761244052742375
w	28	0.030107526881720432	5.05373198393779
d	19	0.02043010752688172	5.613159392551808
A	4	0.004301075268817204	7.861086905995394
h	73	0.07849462365591398	3.6712623471153765
m	25	0.026881720430107527	5.217230716220669
b	5	0.005376344086021506	7.539158811108031
T	1	0.001075268817204301	9.861086905995395
D	3	0.0032258064516129032	8.276124405274238
I	11	0.011827956989247311	6.401655287358096
G	2	0.002150537634408602	8.861086905995395
P	3	0.0032258064516129032	8.276124405274238
M	1	0.001075268817204301	9.861086905995395
g	25	0.026881720430107527	5.217230716220669
l	47	0.05053763440860215	4.306498054317756
O	2	0.002150537634408602	8.861086905995395
H	10	0.010752688172043012	6.539158811108031
N	3	0.0032258064516129032	8.276124405274238
p	5	0.005376344086021506	7.539158811108031
u	3	0.0032258064516129032	8.276124405274238
(4	0.004301075268817204	7.861086905995394
)	4	0.004301075268817204	7.861086905995394
newLine	42	0.04516129032258064	4.468769483216634

Letter	Count	P = Count/N	I= -log2(P)
Number of Letters	751		
Number of Symbols	179		
Number of Characters	930		

$$H(X) = - \sum_{i=1}^{39} p_i \log_2(p_i) = 4.3816093459729935$$

Αρα τελικά έχω εντροπία $H(x) = 4.3816093459729935$

3. Συμπεράσματα

3. Τι συμπεράσματα βγάζετε σχετικά με την εμφάνιση των παραπάνω συμβόλων παρατηρώντας την αυτοπληροφορία τους;

Η πιθανότητα εμφάνισης των γραμμάτων είναι αντιστρόφως αναλογική με την πληροφορία που περιέχει. Όσο μεγαλύτερη είναι η πιθανότητα εμφάνισης τους, τόσο μικρότερη θα είναι η αυτο-πληροφορία που κουβαλάνε.

Είναι πιο κερδοφόρο ένα συχνό γεγονός να παρέχει λιγότερη πληροφορία.

Ερώτημα 2

Να αποδείξετε ότι ισχύει η σχέση $I(X;Y)=I(Y;X)$

assignment-1-2-math

$$I(X;Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(x,y) \cdot \log \left(\frac{P(y|x) \cdot P(x)}{P(x) \cdot P(y)} \right) dx dy$$

$$I(X;Y) = I(Y;X) \iff$$

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(x,y) \cdot \log \left(\frac{P(y|x) \cdot P(x)}{P(x) \cdot P(y)} \right) dx dy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(y,x) \cdot \log \left(\frac{P(x|y) \cdot P(y)}{P(x) \cdot P(y)} \right) dx dy \iff$$

$$\begin{aligned} & \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(x,y) \cdot [\log(P(y|x) \cdot P(x)) - \log(P(x) \cdot P(y))] dx dy = \\ & = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(y,x) \cdot [\log(P(x|y) \cdot P(y)) - \log(P(x) \cdot P(y))] dx dy \iff \end{aligned}$$

$$\begin{aligned} & \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(x,y) \cdot \log(P(y,x)) dx dy - \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(x,y) \cdot \log(P(y) \cdot P(x)) dx dy = \\ & = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(y,x) \cdot \log(P(y,x) \cdot P(y)) dx dy - \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(y,x) \cdot \log(P(y) \cdot P(x)) dx dy \iff \end{aligned}$$

$$\iff 0 = 0$$

Ερώτημα 3

Μία διακριτή πηγή παράγει σύμβολα 0 και 1 και η πιθανότητα εμφάνισης του 0 είναι 0,35.

Τα σύμβολα μεταδίδονται μέσα από ένα κανάλι το οποίο αλλοιώνει το 0 σε 1 με πιθανότητα 0,25 και το 1 σε 0 με πιθανότητα 0,10.

Να υπολογιστούν:

1. Η αυτοπληροφορία κάθε συμβόλου και η εντροπία της πηγής.

1. Η αυτοπληροφορία κάθε συμβόλου και η εντροπία της πηγής.

$$P(x = 0) = 0.35, P(x = 1) = 1 - P(x = 0) = 0.65$$

Εφόσον η πηγή εμφανίζει 0 με πιθανότητα 0.35, τότε αν X η τυχαία τιμή μεταβλητή που είναι εισοδος στο κανάλι, τότε $P(x=0)=0.35$

Επειδή η πηγή παραγει μονο 0 ή 1, τότε ισχυει η σχεση $P(x=1) = 1 - P(x=0) = 0.65$.

Βρισκουμε τις αυτοπληροφοριες του 0 και του 1 :

| Letter | P | I= -log₂(P) |

| ----- | ---- | ----- |

| 0 | 0.35 | 1.515 |

| 1 | 0.65 | 0.622 |

Και την εντροπία :

$$H(X) = - \sum_{i=1}^2 p_i \log_2(p_i) = 0.934$$

$H(x) = 0.934$ bits/symbol

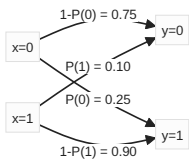
2. Η μέση αμοιβαία πληροφορία μεταξύ εισόδου και εξόδου του καναλιού.

2. Η μέση αμοιβαία πληροφορία μεταξύ εισόδου και εξόδου του καναλιού.

Η μέση αμοιβαία πληροφορία βρίσκεται από τον τύπο :

$$I(X; Y) = \sum_{i=1}^2 \sum_{j=1}^2 P(x_i, y_j) \cdot I(x_i; y_j)$$

Η πιθανότητα αλλοίωσης του 0 στην είσοδο σε 1 στην έξοδο και αντιστοίχα για την αλλοίωση του 1 σε 0, εκφράζει τις εξής δεσμευτικές πιθανότητες αν Y η τ.μ. που περιγράφει την έξοδο του



P(Y/X)	x=0	y=1
x0	0.75	0.25
x1	0.10	0.90

$$P(Y = 1|X = 0) = 0.25 \text{ ενώ } P(Y = 0|X = 1) = 0.1$$

Αντιστοίχα θα έχουμε ότι η πιθανότητα να μην υποστούν αλλοίωση, θα είναι ίση με το συμπλήρωμα της, δηλαδή:

$$P(Y = 1|X = 1) = 1 - P(Y = 0|X = 1) = 0.9 \text{ ενώ } P(Y = 0|X = 0) = 1 - P(Y = 1|X = 0) = 0.25$$

Από τον Νόμο του Bayes, μπορούμε να βρούμε τα $P(Y=y)$ ως : $P(y_j|x_i) = \frac{P(x_i,y_j)}{P(y_j)}$

$$\begin{aligned} P(y = 0) &= P(y = 0|x = 1) * P(x = 1) + \\ &P(y = 0|x = 0) * P(x = 0) \\ &0.1 \cdot 0.65 + 0.75 \cdot 0.35 = \\ &0.328 \end{aligned}$$

$$\begin{aligned} P(y = 1) &= P(y = 1|x = 1) * P(x = 1) + \\ &P(y = 1|x = 0) * P(x = 0) \\ &0.9 \cdot 0.65 + 0.25 \cdot 0.35 = \\ &0.673 \end{aligned}$$

Οι τιμές όλων των ενδεχομένων :

P	Value
P(X=0)	0.35
P(X=1)	0.65
P(Y=0)	0.328
P(Y=1)	0.773
P(Y=0 X=0)	0.75
P(Y=1 X=0)	0.25
P(Y=0 X=1)	0.1
P(Y=1 X=1)	0.90

Από τον Νόμο του Bayes το $I(X;Y)$ γίνεται :

$$I(X; Y) = \sum_{i=1}^2 \sum_{j=1}^2 P(y_j|x_i) \cdot P(x_i) \cdot I(x_i; y_j)$$

Επιπλέον έχουμε ότι $I(x_i; y_j) = I(y_j; x_i) = \log_2 \left(\frac{P(y_j|x_i)}{P(y_j)} \right)$

$$I(y = 0; x = 0) = \log_2 \left(\frac{P(y = 0|x = 0)}{P(y = 0)} \right) = \log_2 \left(\frac{0.75}{0.328} \right) = 1.193$$

$$I(y = 1; x = 0) = \log_2 \left(\frac{P(y = 1|x = 0)}{P(y = 1)} \right) = \log_2 \left(\frac{0.25}{0.672} \right) = -1.427$$

$$I(y = 0; x = 1) = \log_2 \left(\frac{P(y = 0|x = 1)}{P(y = 0)} \right) = \log_2 \left(\frac{0.1}{0.328} \right) = -1.714$$

$$I(y = 1; x = 1) = \log_2 \left(\frac{P(y = 1|x = 1)}{P(y = 1)} \right) = \log_2 \left(\frac{0.9}{0.672} \right) = 0.421$$

Οπότε έχουμε:

$$\begin{aligned}
I(X;Y) &= P(y=0|x=0) * P(x=0) * I(y=0; x=0) + \\
&\quad P(y=1|x=0) * P(x=0) * I(y=1; x=0) + \\
&\quad P(y=0|x=1) * P(x=1) * I(y=0; x=1) + \\
&\quad P(y=1|x=1) * P(x=1) * I(y=1; x=1) = \\
&\quad (0.75) * (0.35) * (1.193) + \\
&\quad (0.25) * (0.35) * (-1.427) + \\
&\quad (0.1) * (0.65) * (-1.714) + \\
&\quad (0.9) * (0.65) * (0.421) = \\
&\quad 0.323
\end{aligned}$$

Η μέση αμοιβαία πληροφορία μεταξύ εισόδου και εξόδου του καναλιού είναι : $I(X;Y) = 0.323$