

1)

Прописываем кол-во символов в строке

2)

Сортируем символы так по частоте упоминания и по алфавиту (пробел идет раньше букв)

ПРОБЕЛ	57
о	41
е	40
т	37
с	33
и	28
р	26
в	20
а	19
м	17
н	17
я	10
п	8
у	7
ы	7
б	6
д	6
з	6
ч	6
г	4
к	4
ф	4
щ	4
ь	4
л	2
ц	2
ю	2
ё	1
х	1

3) считаем кол-во символов СУММ

4) делим кол-во на 2

5) проходимся по каждому символу так:

пока значение символов предыдущих в группе + значение текущего < максимума (той суммы что делили на 2), то это группа [0], как только сумма перевалит максимум, то это группа 1.

После этого у нас появляется 2 подгруппы следующего уровня. Для них мы вычисляем максимум - сумма всех в группе/2, округленное к меньшему целому. После этого делаем с этой группой пункт 5 до тех пор, пока у каждого символа в группе не будет своего значения

(В общем сверху всегда 0, главное найти точку перехода. ЕСЛИ ЗНАЧЕНИЕ ПОПАДАЕТ РОВНО В ТОЧКУ ПЕРЕХОДА, то мы его относим к 1 а не к 0)  
 6) И потом отдельным столбцом можно записать сами коды (1 и 0 из строки)

1	ПРОБЕЛ	9	0	0					0
2	и	8	0	1	0				10
3	е	5	0	1	1				11
4	а	3	1	0	0	0			1000
5	в	3	1	0	0	1			1001
6	м	3	1	0	1	0			1010
7	л	2	1	0	1	1	0		10110
8	о	2	1	0	1	1	1		10111
9	п	2	1	1	0	0			1100
10	р	2	1	1	0	1	0		11010
11	т	2	1	1	0	1	1		11011
12	х	2	1	1	1	0	0		11100
13	д	1	1	1	1	0	1		11101
14	к	1	1	1	1	1	0	0	111100
15	н	1	1	1	1	1	0	1	111101
16	с	1	1	1	1	1	0	0	111110
17	ч	1	1	1	1	1	1	0	1111110
18	ы	1	1	1	1	1	1	1	1111111
19		49	24,5	13	8	5			
20			22	6,5	4	2,5			
21			11						
22									
23			27	14					
24			13,5	7					

7)  $=-1*B2/ \$B\$19*LOG(B2/ \$B\$19;2)$  для каждого считаем по этой формуле, где  $\$B\$19$  это сумма всех символов, и B2 - текущий символ

$=-1*B1/ \$B\$19*LOG(B1/ \$B\$19;2)$									
0						00	0		
1	0					010			
	1					011			

ЕСЛИ <span>✖</span> <span>✓</span> $f_x$ $=-1*B1/\$B\$19*LOG(B1/\$B\$19;2)$												
	A	B	C	D	E	F	G	H	I	J	K	L
1	ПРОБЕЛ	9	0	0						0		$B1/\$B\$19;2$
2	и	8	0	1	0					10		0,4268914
3	е	5	0	1	1					11		0,33599814
4	а	3	1	0	0	0				1000		0,24671923
5	в	3		0	0	1				1001		0,24671923
6	м	3		0	1	0				1010		0,24671923
7	л	2		0	1	1	0			10110		0,1883555
8	о	2		0	1	1	1			10111		0,1883555
9	п	2		1	0	0				1100		0,1883555
10	р	2		1	0	1	0			11010		0,1883555
11	т	2		1	0	1	1			11011		0,1883555
12	х	2		1	1	0	0			11100		0,1883555
13	д	1		1	1	0	1			11101		0,11458592
14	к	1	1	1	1	1	0	0	111100		0,11458592	
15	н	1	1	1	1	1	0	1	111101		0,11458592	
16	с	1	1	1	1	1	1	0	111110		0,11458592	
17	ч	1	1	1	1	1	1	1	1111110		0,11458592	
18	ы	1	1	1	1	1	1	1	1111111		0,11458592	
19		49	24,5	13	8	5						
20			22	6,5	4	2,5						
21			11									
22												
23			27	14								
24			13,5	7								

8)Энтропия - сумма вычисленной колонки:

ЕСЛИ <span>⬆</span> <span>✖</span> <span>✓</span> $f_x$ =СУММ(L1:L18)												
	A	B	C	D	E	F	G	H	I	J	K	L
1	ПРОБЕЛ	9	0	0						0		0,44904211
2	и	8	0	1	0					10		0,4268914
3	е	5	0	1	1					11		0,33599814
4	а	3	1	0	0	0				1000		0,24671923
5	в	3	1	0	0	1				1001		0,24671923
6	м	3	1	0	1	0				1010		0,24671923
7	л	2	1	0	1	1	0			10110		0,1883555
8	о	2	1	0	1	1	1			10111		0,1883555
9	п	2	1	1	0	0				1100		0,1883555
10	р	2	1	1	0	1	0			11010		0,1883555
11	т	2	1	1	0	1	1			11011		0,1883555
12	х	2	1	1	1	0	0			11100		0,1883555
13	д	1	1	1	1	0	1			11101		0,11458592
14	к	1	1	1	1	1	0	0		111100		0,11458592
15	н	1	1	1	1	1	0	1		111101		0,11458592
16	с	1	1	1	1	1	1	0		111110		0,11458592
17	ч	1	1	1	1	1	1	1	0	1111110		0,11458592
18	ы	1	1	1	1	1	1	1	1	1111111		0,11458592
19		49	24,5	13	8	5					Энтропия:	L18)
20			22	6,5	4	2,5						
21			11									
22												
23			27	14								
24			13,5	7								

9)Стоимость дерева  
 $=B1/\$B\$19*ДЛСТР(J1)$

где B1 - символ,  $\$B\$19$  - общая сумма символов, ДЛСТР(J1) - длинна кода символа (на скрине не 00 а 00 т.к. У меня не хочет 2 ноля ставить превращает в 1)

ЕСЛИ <span>✖</span> <span>✓</span> $f_x$ =B1/\$B\$19*ДЛСТР(J1)															
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
1	ПРОБЕЛ	9	0	0						∞		0,44904211		ДЛСТР(J1)	
2	и	8	0	1	0					10		0,4268914		0,32653061	
3	е	5	0	1	1					11		0,33599814		0,20408163	
4	а	3	1	1	0	0	0			1000		0,24671923		0,24489796	
5	в	3		1	0	0	1			1001		0,24671923		0,24489796	
6	м	3		1	0	1	0			1010		0,24671923		0,24489796	
7	л	2		1	0	1	1	0		10110		0,1883555		0,20408163	
8	о	2		1	0	1	1	1		10111		0,1883555		0,20408163	
9	п	2		1	1	0	0	0		1100		0,1883555		0,16326531	
10	р	2			1	1	0	1	0		11010		0,1883555		0,20408163
11	т	2			1	1	0	1	1		11011		0,1883555		0,20408163
12	х	2			1	1	0	0	0		11100		0,1883555		0,20408163
13	д	1		1	1	1	0	1		11101		0,11458592		0,10204082	
14	к	1	1	1	1	1	0	0	111100		0,11458592		0,12244898		
15	н	1	1	1	1	1	0	1	111101		0,11458592		0,12244898		
16	с	1	1	1	1	1	1	0	111110		0,11458592		0,12244898		
17	ч	1	1	1	1	1	1	1	0	1111110		0,11458592		0,14285714	
18	ы	1	1	1	1	1	1	1	1	1111111		0,11458592		0,14285714	
19		49	24,5	13	8	5					Энтропия:	3,76973784			
20			22	6,5	4	2,5									
21			11												
22															
23			27	14											
24			13,5	7											
25															

ЕСЛИ <span>✖</span> <span>✓</span> $f_x$ =СУММ(N1:N18)														
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	ПРОБЕЛ	9	0	0						∞		0,44904211		0,36734694
2	и	8	0	1	0					10		0,4268914		0,32653061
3	е	5	0	1	1					11		0,33599814		0,20408163
4	а	3	1	0	0	0				1000		0,24671923		0,24489796
5	в	3	1	0	0	1				1001		0,24671923		0,24489796
6	м	3	1	0	1	0				1010		0,24671923		0,24489796
7	л	2	1	0	1	1	0			10110		0,1883555		0,20408163
8	о	2	1	0	1	1	1	1		10111		0,1883555		0,20408163
9	п	2	1	1	0	0				1100		0,1883555		0,16326531
10	р	2	1	1	0	1	0			11010		0,1883555		0,20408163
11	т	2	1	1	0	1	1			11011		0,1883555		0,20408163
12	х	2	1	1	1	0	0			11100		0,1883555		0,20408163
13	д	1	1	1	1	0	1			11101		0,11458592		0,10204082
14	к	1	1	1	1	1	0	0		111100		0,11458592		0,12244898
15	н	1	1	1	1	1	0	1		111101		0,11458592		0,12244898
16	с	1	1	1	1	1	0	0		111110		0,11458592		0,12244898
17	ч	1	1	1	1	1	1	1	0	1111110		0,11458592		0,14285714
18	ы	1	1	1	1	1	1	1	1	1111111		0,11458592		0,14285714
19		49	24,5	13	8	5					Энтропия:	3,76973784	стоим. Древа	N18)
20			22	6,5	4	2,5								
21			11											
22														
23			27	14										
24			13,5	7										

Ну и стоимость - сумма колонки

10) кол-во информации на 1 двоичный символ



	P	Q
		18,9994787
		16,8884255
		10,555266
		12,6663192
		12,6663192
		12,6663192
		10,555266
		10,555266
		8,44421277
		10,555266
		10,555266
		10,555266
		5,27763298
		6,33315958
		6,33315958
		6,33315958
		7,38868617
		7,38868617
символ	1,0555266	Q18)