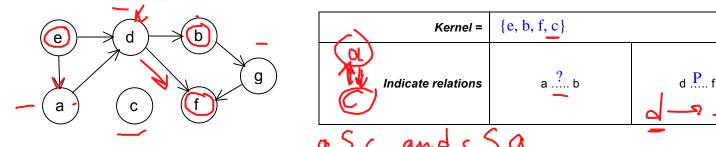
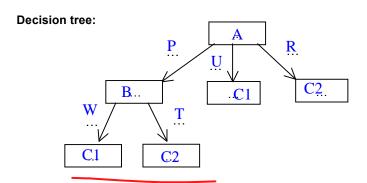
First and last name:	Sofya Akse	enyuk						Index:	150)284	1	Sco	ore		/ 20p	
1. [3p] Given the represent group these users into clu- Also, a similarity matrix betale.g., G1: U2, U5). Compu 2nd iteration based on the	sters. One has tween the user te the value of	already s and th the J m	drawn e centr	the thoids ha	ree cer as been	ntroids (C1-C3 ited. Pr	; e.g., (esent t	C1 is a he gro	cen ups (troid o	of grou 3) obta	p G1) i ined af	n the ter th	1st ite e 1st it	ration. eraion
Representation of 7 users (objects) Similarity matrix between the centroids and the users: Clustering in the 1 st iterat.													at.			
in terms of 2 documents (features): U1 U2 U3 U4 U5 U6 U7 for each cluster, list the use												sers				
		U7	C1	0.83	0.95	(J)	0.92	0.87	0.99	0.9	94	G1: [13, U6			
	0.2 0.8 0.6 0.6 0.3 0.8	0.2	C2	0.61	0.99	0.94	0.99	0.67	0.96	1	0)	_	2, U4		7	
		0.5	C3 (1.0	0.63	0.83	0.56	0.99	0.79	0.6	31	G3: U	1, U:	5_		
Centroids in the											L			<u> </u>		
C1 C2			Lmos	euro s	after th	o 1 st ita	ration									
	0,73				9) + (-		1) + (1	+0.9	9) =	6 96					
D2 0,75 0,6	0,23		J – (110,7	<i>(</i>)	0,221	J•77 1 .	1) (1	110.0	<i>)</i>	0,70					
2. [2.5p] Given the classi 6-NN (6-Nearest Neighbo determine the class of Y. C	ur) classifier i	ncorpora	ting the	e major	rity rule	voting	or the v	weighte	d votir	ıg. Li	st the	docun	nents th	at yo	u cons	
D1	1 D2	D3		D4	D5		D6	D7		D8		D9	D.	10	Y	
Class B		В		Α	В		Α	Α		В		С	(?	
Similiraity with Y 0.7	7 0.8	0.9	(0.65	1.0		0.6	0.7	5	0.5		8.0	0.	9	-	
Documets considered to classify Y in 6-NN:																
Documets considered	to classify i	III O-ININ.														
6-NN (majority rule)	Class A:					Class B	: 3					Class	C:	1		
	Class A:			d to cla			: 3					Class	C:	1		
	Class A:	2 er: Y is a			ass: E			'+0.9+	1=2.6	5		Class		.8		
6-NN (majority rule)	Class A: Answ Class A:	2 er: Y is a	assigne 5+0,6=	=1.25	ass: E	3		′+0.9+	1=2.6	5						
6-NN (majority rule)	Class A: Answe	2 er: Y is a 0,65 er: Y is a	assigne 5+0,6=	=1.25 d to cla	ass: E	B	: 0.7	/+0.9+	1=2.6	5		Class		.8		
6-NN (majority rule) 6-NN (weighted voting) 3. [1.5p] Given the confuthe cost matrix for the class	Class A: Answell Class A: Answell Answell Assignment ix selfication problem.	2 er: Y is a 0,65 er: Y is a and olem	assigne 5+0,6= assigne	=1.25 d to cla	ass: E	B	: 0.7	C5	1=2.6	5	C1	Class	C: 0	.8	C4	C5
6-NN (majority rule) 6-NN (weighted voting) 3. [1.5p] Given the confuthe cost matrix for the clasinvolving five classes C1-Cin rows; predicted classes	Class A: Answer Class A: Answer Answer Assignment ix selfication problem in columns)	2 er: Y is a 0,65 er: Y is a and olem oses and	assigne 5+0,6= assigne	=1.25 d to cla Cor C1 10	ass: E	Class B matrix C3 2	: 0.7 (C4 0	C5 0	C	1	0	Cos C2 3	C: 0 t matrix C3 4	.8	1	2
6-NN (majority rule) 6-NN (weighted voting) 3. [1.5p] Given the confuthe cost matrix for the clasinvolving five classes C1-C in rows; predicted classes 100 documents, compute	Class A: Answer Class A: Answer An	2 er: Y is a 0,65 er: Y is a and olem olem olem olem olem olem olem olem	assigne 5+0,6= assigne	=1.25 d to cla Cor C1 10 2 •	nfusion C2 0 16	B Class B Clas	: 0.7 C4 0 8	C5 0 0	C	1 2	0 1 7	Cos	C: 0	.8	1 1 &	2
6-NN (majority rule) 6-NN (weighted voting) 3. [1.5p] Given the confuthe cost matrix for the clasinvolving five classes C1-Cin rows; predicted classes	Class A: Answer Class A: Answer An	2 er: Y is a 0,65 er: Y is a and olem osses and dition ossifi-	assigne 5+0,6= assigne C1 C2 C3	=1.25 d to cla Cor C1 10 2 • • •	ass: E	Class B matrix C3 2 0 14	C4 0 8 • 0	C5 0	C	1 2 3	0 1 7 1 7	Cos	C: 0	.8	1	2
6-NN (majority rule) 6-NN (weighted voting) 3. [1.5p] Given the confuthe cost matrix for the clasinvolving five classes C1-Cin rows; predicted classes 100 documents, compute accuracy, recall for class C	Class A: Answer Class A: Answer An	er: Y is a 0,65 er: Y is a and olem sses and otion ssifi-	assigne 5+0,6= assigne	=1.25 d to cla Cor C1 10 2 •	nfusion C2 0 16 4	Class B B Class B C3 C3 C3 C4	: 0.7 C4 0 8	C5 0 0	C	1 2 3 4	0 1 7	Cos C2 3 0 1 1 1 0 . 0	C: 0 t matrix	.8	1 1 % 1	2 2 3
6-NN (majority rule) 6-NN (weighted voting) 3. [1.5p] Given the confuthe cost matrix for the clasinvolving five classes C1-Cin rows; predicted classes 100 documents, compute accuracy, recall for class C	Class A: Answer Class A: Answer An	2 er: Y is a 0,65 er: Y is a and olem olem olem olem olem olem olem olem	assigne 5+0,6= assigne C1 C2 C3 C4 C5	Cor C1 10 2 · 2 · 4 · 4 · 4	nfusion C2 0 16 4 0 0	Class B matrix C3 2 0 14 4 0	C4 0 8 • 0 10	C5 0 0 0 0		1 2 3 4	0 1 7 1 7 1 7	Cos C2 3 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	C: 0	.8	1 1 % 1 0	2 2 3 1
6-NN (majority rule) 6-NN (weighted voting) 3. [1.5p] Given the confuthe cost matrix for the clasinvolving five classes C1-C in rows; predicted classes 100 documents, compute accuracy, recall for class Cation cost	Class A: Answ. Class A: Answ. Answ. Sification probes in columns) the classification misclassification probes. (10+16+14+14)	2 er: Y is a 0,65 er: Y is a and olem osses and olicition ossifi- 10+20	assigne 5+0,6= assigne C1 C2 C3 C4 C5)/100	Col C1 10 2 • • • • • • • • • • • • • • • • • • •	nfusion C2 0 16 4 0 0 100	Class B matrix C3 2 0 14 4 0 Recall	C4 0 8 0 10 2 for C4	C5 0 0 0 0 20 = 10	C C C C C C C C C C C C C C C C C C C	1 2 3 4 5 2+4	0 1 7 1 7 1 2 1 4) = 1	Class Cos C2 3 0 11 0 10 0/16	C: 0 t matrix	.8	1 8 1 0 1 2	2 2 3 1
6-NN (majority rule) 6-NN (weighted voting) 3. [1.5p] Given the confuthe cost matrix for the clasinvolving five classes C1-Cin rows; predicted classes 100 documents, compute accuracy, recall for class Cation cost	Class A: Answer Class A: Answer An	2 er: Y is a 0,65 er: Y is a and olem sses and dition ssifi- -10+20 3*1 + 1 5 3 7 2 4 er to X	assigne 5+0,6= assigne C1 C2 C3 C4 C5)/100 *2 + 4	Col C1 10 2	nfusion C2 0 16 4 0 0 100 2*1 + 4 7 5 3 rbitrary	Class B matrix C3 2 0 14 4 0 Recall 4*5 +	C4 0 8 0 10 10 2 for C4 4*1 + represented the representation of th	C5 0 0 0 0 20 = 10 - 2*1 =	C C C C C C C C C C C C C C C C C C C	11 22 33 44 55 22+4 +8+	0 1 1 1 1 1 1 1 1 1 1	Class Cos C2 3 0 1 1 0/16 +2+20 esent a	C: 0 t matrix	= 52 of c	1 1 8 1 0 1 2 hromos	2 2 3 1 0
6-NN (majority rule) 6-NN (weighted voting) 3. [1.5p] Given the confuthe cost matrix for the clasinvolving five classes C1-Cin rows; predicted classes 100 documents, compute accuracy, recall for class Cation cost Classification accuracy = Misclassification cost = 2° 4. [2p] Given two chromosobtained after applying or	Class A: Answ. Class A: Answ. Class A: Answ. Class A: Answ. Signature Coriginal class in columns) the classification problem in columns the classification in columns Coriginal class in columns Coriginal class in columns Coriginal class Coriginal	2 er: Y is a 0,63 er: Y is a and olem osses and olem ossifi- 10+20 3*1 + 1 5 3 7 2 4 ver to X pplying of a)	21 C2 C3 C4 C5)/100 *2 + 4 6] and Y cycle c	Col C1 10 2 4 1+1 + d Y = [with ar rossov	nfusior C2 0 16 4 0 0 100 2*1+ 4 7 5 3 rbitrary ver to X	Class B R C3 2 0 14 4 7 Recall 4*5 + 2 6 1] parts irr and Y	C4 0 8 0 10 10 2 for C4 4*1 + represented the representation of th	$ \begin{array}{c} $	C C C C C C C C C C C C C C C C C C C	11 22 33 44 55 22+4 +8+6 attion	0 1 7 1 7 1 2 1 4 1 - 2 1 - 4 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 -	Class Cos C2 3 0 1 1 0/16 +2+20 esent a	C: 0 t matrix	= 52 of c	1 1 8 1 0 1 2 2 hromos	2 2 3 1 0
6-NN (majority rule) 6-NN (weighted voting) 3. [1.5p] Given the confuthe cost matrix for the clasinvolving five classes C1-Cin rows; predicted classes 100 documents, compute accuracy, recall for class Cation cost Classification accuracy = Misclassification cost = 2. 4. [2p] Given two chromosobtained after applying or b) a pair of chromosome of	Class A: Answer Class A: Answer Class A: Answer Answer Signature Continuous classification probes Continuous classification probes Continuous classification Class A: Answer Answer Class A: Answer Class A: Answer Class A: Answer Class A: Answer Continuous classification Continuous classification Class A: Answer Answer Answer Class A: Answer Answer Answer Class A: Answer Answer	2 er: Y is a 0,63 er: Y is a and olem sses and dition ssifi- 0.3*1 + 1 5 3 7 2 4 er to X pplying of a) 5 3	21	Col C1 10 2	nfusior C2 0 16 4 0 0 100 2*1 + 4 7 5 3 rbitrary ver to X	Class B R C3 2 0 14 4 7 Recall 4*5 + 2 6 1] parts irr and Y	C4 0 8 0 10 10 2 for C4 4*1 + represented	$ \begin{array}{c} $	C C C C C C C C C C C C C C C c c c c c	11 22 33 44 55 22+4 +8+6 attion	0 1 7 1 7 1 2 1 4 1 - 2 1 - 4 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 -	Class Cos C2 3 0 1 1 0/16 +2+20 essent alren be	C: 0 t matrix	= 52 $= 52$ of cof cof 22 nd and 3	1	2 2 3 1 0
6-NN (majority rule) 6-NN (weighted voting) 3. [1.5p] Given the confuthe cost matrix for the clasinvolving five classes C1-Cin rows; predicted classes 100 documents, compute accuracy, recall for class Cation cost Classification accuracy = Misclassification cost = 2. 4. [2p] Given two chromosobtained after applying or b) a pair of chromosome of a) order 1 crossover (pair confusion cost in the cost in t	Class A: Answer Class A: Answer Class A: Answer Answer Signature Continuous classification problem Continuous cl	2 er: Y is a 0,63 er: Y is a and olem osses and olem ossifi- 10+20 3*1 + 1 5 3 7 2 4 ver to X pplying of a)	21	Col C1 10 2	nfusior C2 0 16 4 0 0 100 2*1 + 4 7 5 3 rbitrary ver to X	Class B R matrix C3 2 0 14 4 0 Recall 4*5 + 6 2 6 1] parts ir and Y. b) (C4 0 8 0 10 2 for C4 4*1 + repressible rited	C5 0 0 0 20 = 10 - 2*1 =	C C C C C C C C C C C C C C C C C C C	11 22 33 44 55 22+4 +8+ eation rective	0 17 17 12 14) = 1 -2+4- ns, prese child	Class Cos C2 3 0 1 1 0/16 +2+20 essent alren be	C: 0 t matri:	= 52 $= 52$ of cof cof 22 nd and 3	1	2 2 3 1 0

5. [2p] Given the outranking graph concerning seven alternatives a-g, find its **kernel** (you do not need to present the steps of the formal algorithm involving the analysis of predecessors) and indicate the relations holding for the two showed pairs of alternatives from among P (preference), I (indifference), and ? (incomparability). An incorrect answer cancels out the points for one correct answer.



6. [3.5p] Given the information table referring to condition attributes A and B and a decision attribute C (class C1 or C2), compute Ent(C), Ent(C,A), Ent(C,B), InformationGain(C,A) and InformationGain(C,B) in the first iteration. Fill in the entire decision tree obtained with **the ID3 algorithm** incorporating information gain for splitting in each node. You do not need to present any computations for the second tree level (if needed) - since there are only two attributes, it is natural that in the second iteration you will use the attribute that has not been used in the first iteration. Remember that: $log_21 = 0$, $log_2(1/2) = -1$, $log_2(1/3) = -1.585$, $log_2(2/3) = -0.585$, $0 \cdot log_20 = 0$; $1/2log_2(1/2) + 1/2log_2(1/2) = -1$; $1/3log_2(1/3) + 2/3log_2(2/3) = -0.918$; 3/8*0.918 = 0.344.

	4			4/041 0 4/0 4/041 0 4/0 0 5 1 0 5 1
	Α	В	C	$Ent(C) = \frac{4/8*\log 2}{4/8} + \frac{4/8*\log 2}{4/8} = 0.5 + 0.5 = 1$
1	• U	Т	C1\	
2	U •	I	C1	$Ent(C,A) = \frac{4/8}{8}(-2/4*\log 2 2/4 - 2/4*\log 2 2/4 - 0/4*\log 2 0/4) + 4/8*(-0/4*\log 2 0/4 - 1/4*\log 2 1/4 - 0/4*\log 2 0/4) + 4/8*(-0/4*\log 2 0/4 - 1/4*\log 2 1/4 - 0/4*\log 2 0/4) + 4/8*(-0/4*\log 2 0/4 - 1/4*\log 2 1/4 - 0/4*\log 2 0/4) + 4/8*(-0/4*\log 2 0/4 - 1/4*\log 2 1/4 - 0/4*\log 2 0/4) + 4/8*(-0/4*\log 2 0/4 - 1/4*\log 2 1/4 - 0/4*\log 2 0/4) + 4/8*(-0/4*\log 2 0/4 - 1/4*\log 2 1/4 - 0/4*\log 2 0/4) + 4/8*(-0/4*\log 2 0/4 - 1/4*\log 2 1/4 - 0/4*\log 2 0/4) + 4/8*(-0/4*\log 2 0/4 - 1/4*\log 2 1/4 - 0/4*\log 2 0/4) + 4/8*(-0/4*\log 2 0/4 - 1/4*\log 2 1/4 - 0/4*\log 2 0/4) + 4/8*(-0/4*\log 2 0/4 - 1/4*\log 2 1/4 - 0/4*\log 2 0/4) + 4/8*(-0/4*\log 2 0/4 - 1/4*\log 2 1/4 - 0/4*\log 2 0/4) + 4/8*(-0/4*\log 2 0/4 - 1/4*\log 2 1/4 - 0/4*\log 2 0/4) + 4/8*(-0/4*\log 2 0/4 - 1/4*\log 2 1/4 - 0/4*\log 2 0/4) + 4/8*(-0/4*\log 2 0/4 - 1/4*\log 2 1/4 - 0/4*\log 2 0/4) + 4/8*(-0/4*\log 2 0/4 - 1/4*\log 2 1/4 - 0/4*\log 2 0/4) + 4/8*(-0/4*\log 2 0/4 - 1/4*\log 2 0/4) + 4/8*(-0/4*\log 2 0/4) + 4/8*(-0/4*(-0/4*\log 2 0/4)) + 4/8*(-0/4*(-0/$
3	• P	W	¢ 1	$-3/4*\log 2 \ 3/4) = 0.5*(0.5+0.5+0) + 0.5*(0+0.5+0.31125) = 0.5+0.4=0.9$
4 -	P.	–W	CT	$Ent(C,B) = 4/8*(-2/4*log2\ 2/4 - 2/4*log2\ 2/4) + 4/8*(-2/4*log2\ 2/4 - 2/4*log2\ 2/4) = 1$
5_	P	T/	C2 \	1 00 01
6	R•	D	C2	InformationGain(C,A) = Ent(C) - Ent(C,A) = $1 - 0.9 = 0.1$
7	R	W	C2	1 1 - 0
8	R	W	C2/	InformationGain(C,B) = Ent(C) - Ent(C,B) = $1-1=0$



7. [2.5p] Given the 4x4 matrix of inputs (see below), **convolve** it with the 3x3 filter (**the bias value is given below the 3x3 matrix**) with a stride of 1, then apply **ReLU** on the matrix obtained after the convolution, and finally apply **average (AVE) pooling** with a filter of **size 2x2** and a stride of 2 on the matrix obtained after ReLU. You need to determine the size of matrices obtained after each operation on your own (the 4x5 matrices given below serve only for filling the results, but you should use only as many as cells as you need).

input (4x4) filter (3x3)			after convolution						after ReLU					_	after AVE pooling								
0	-2	0	-1	1	0	0	-2	-3	2	-1	0												
0	2	1	-2	0	-1	0																	
-2	1	0	0	0	0	1																	
1	0	2	1	bi	as =	+2																	

8. [3p] Find the shortest path from **s** to **t** in the graph presented below using the **A*** (A star) algorithm (the distances between the nodes are given next to the edges and the heuristic distances **h** from node **t** are given in bold next to the nodes). Fill in the table showing the process of finding the shortest pass. Indicate the shortest path and provide its length. To denote the visited node in the first column, use the "X" symbol. You do not need to cross out the distances or previous nodes that are updated in the next iterations, but they need to be visible in the table, so do not delete them either.

15 S 9 5	Visit?	Node	Shortest distance from s	Heuristic distance to t	Total distance f = g + h	Previous node
10 a 2 b 11	X	s	0	15	15	
3		а	6 9	10	19	S
5 6 7	X	b	\$ 5	11	16	S
5 C d 5	X	C	† 11	5	16	b
6 5		d	№ 18	5	23	c
t	X	t	% 17	0	17	c

Answer: The shortest path from **s** to **t** is: s - b - c - t

The length of the shortest path is: 48