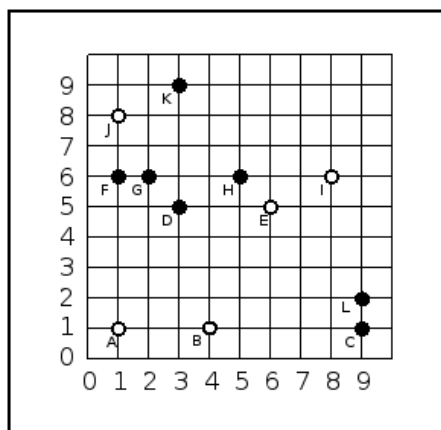


Άσκηση εμπέδωσης X Fold-Cross-Validation, Accuracy, Precision, Recall

Έστω το παρακάτω σύνολο δεδομένων.



i) Να δημιουργήσετε τα σύνολα εκπαίδευσης (training) και δοκιμής (testing) κατάλληλα για 3-fold cross validation εξετάζοντας τα δεδομένα με αλφαβητική σειρά. Στην απάντηση θα πρέπει να σημειώσετε ποιο στιγμιότυπο συμπεριλαμβάνεται σε ποιο σύνολο

F1: Train Set:	A B C D E F G H	Test Set:	I J K L
F2: Train Set:	A B C D I J K L	Test Set:	E F G H
F3: Train Set:	I J K L E F G H	Test Set:	A B C D

ii) Χρησιμοποιήστε τον αλγόριθμο κατηγοριοποίησης του ενός εγγύτερου γείτονα υιοθετώντας ως μέτρο ομοιότητας την Ευκλείδεια απόσταση να υπολογίσετε την ακρίβεια (Accuracy), Ορθότητα (Precision) και Ευαισθησία (Recall) με βάση τη μέθοδο επικύρωσης 3-fold cross validation στα σύνολα δεδομένων που δημιουργήσατε στο προηγούμενο ερώτημα

	Πρόβλεψη	
	P	N
P	TP	FN
N	FP	TN

F1:

	Πρόβλεψη	
	P	N
P	1	1
N	0	2

ACCURACY: $(TP+TN) / (TP+FP+FN+TN) = 3/4$

PRECISION: $TP / (TP+FP) = 1/1$

RECALL: $TP / (TP+FN) = 1/2$

F2:

	Πρόβλεψη	
	P	N
P	1	0
N	1	2

ACCURACY: $(TP+TN) / (TP+FP+FN+TN) = 3/4$

PRECISION: $TP / (TP+FP) = 1/2$

RECALL: $TP / (TP+FN) = 1/1$

F3:

	Πρόβλεψη	
	P	N
P	1	1
N	0	2

ACCURACY: $(TP+TN) / (TP+FP+FN+TN) = 3/4$

PRECISION: $TP / (TP+FP) = 1/1$

RECALL: $TP / (TP+FN) = 1/2$

Accuracy = $(0.75 + 0.75 + 0.75)/3 = 0.75$

Precision = $(1 + 0.5 + 1) / 3 = 0.83$

Recall = $(0.5+1+0.5)/3 = 0.67$