

**Τζόγκα Κώστας 1115201600171**

**Βαγγέλης Νάζαρης 1115201600114**

### **Εκτέλεση**

**A: `python3 predict.py -i <input file> -a <actual file> -c <config file>`**

**B: `python3 new_representation.py -i <input file> -c <config file>`**

**Γ: `python3 convert_csv_file.py -i <input file> -o <output file>`**

A: Η εύρεση των σφαλμάτων **mae** και **mse** γίνεται με χρήση της συνάρτησης `evaluate`. Το σφάλμα **mape** υπολογίζεται έπειτα από αντικατάσταση των `actual` μηδενικών τιμών με τον μέσο όρο της γραμμής της οποίας βρίσκεται.

B: Το νέο νευρωνικό δημιουργείται με χρήση της συνάρτησης `Model`. Η έξοδος του οποίου είναι τα διανύσματα 64 συντεταγμένων.

Γ: Με αυτό το πρόγραμμα γίνεται η μετατροπή των αρχείων με τα διανύσματα (**`new_representations.csv`**, **`nn_representations.csv`**) σε μορφή κατανοητή για το πρόγραμμα `clustering` της 2ης εργασίας. Αυτό ουσιαστικά γίνεται αντικαθιστώντας τα κόμματα (',') που χωρίζουν τις συντεταγμένες με κενό (' ') κι ενώνοντας την ημερομηνία με την ώρα με το σύμβολο & ώστε το πρόγραμμα `clustering` να το εκλάβει ως το όνομα του διανύσματος. Πχ το `2016-04-01 05:00:00 0.1, 0.2` μετατρέπεται σε `2016-04-01&05:00:00 0.1 0.2`.

# Clustering

## 4 Εποχές

new representation	Silhouette	time
Cluster.1	0.622834	
Cluster.2	0.285999	
Cluster.3	0.294095	
Cluster4	0.335154	
Alg:222	0.38452	0.115079

Algorithm: I2A2U2 - Silhouette: [0.622834, 0.285999, 0.294095, 0.335154, 0.38452]

Algorithm: I2A2U1 - Silhouette: [0.325714, 0.603129, 0.21274, 0.351525, 0.373277]

Algorithm: I2A1U2 - Silhouette: [0.267316, 0.304207, 0.279239, 0.659169, 0.37748]

Algorithm: I2A1U1 - Silhouette: [0.313019, 0.225849, 0.618508, 0.16765, 0.331256]

Algorithm: I1A2U2 - Silhouette: [0.264702, 0.3018, 0.281818, 0.659882, 0.377051]

Algorithm: I1A2U1 - Silhouette: [0.178714, 0.264821, 0.289359, 0.632602, 0.34137]

Algorithm: I1A1U2 - Silhouette: [0.661225, 0.2796, 0.302795, 0.26877, 0.378097]

Algorithm: I1A1U1 - Silhouette: [0.32916, 0.652478, 0.252284, 0.295279, 0.3823]

### Παρατηρήσεις:

- Το Initialization 1 και 2 και το Assignment 1 και 2 έχουν περίπου ίδια απόδοση. Ενώ το Update 2 έχει καλύτερη απόδοση απο το Update 1.

nn representation	Silhouette	time
Cluster.1	0.235934	
Cluster.2	0.269775	
Cluster.3	0.303483	
Cluster4	0.345799	
Alg:212	0.288748	0.49559

Algorithm: I2A2U2 - Silhouette: [0.31915, 0.255529, 0.232245, 0.214778, 0.255426]

Algorithm: I2A2U1 - Silhouette: [0.256339, 0.177478, 0.338459, 0.321519,0.273449]

Algorithm: I2A1U2 - Silhouette: [0.235934, 0.269775, 0.303483, 0.345799,0.288748]

Algorithm: I2A1U1 - Silhouette: [0.25776, 0.333684, 0.276488, 0.248588, 0.27913]

Algorithm: I1A2U2 - Silhouette: [0.223013, 0.23535, 0.207515, 0.27364, 0.23488]

Algorithm: I1A2U1 - Silhouette: [0.250177, 0.228636, 0.188765, 0.209303, 0.21922]

Algorithm: I1A1U2 - Silhouette: [0.208634, 0.226563, 0.221851, 0.256829,0.228469]

Algorithm: I1A1U1 - Silhouette: [0.139379, 0.186497, 0.215005, 0.222032,0.190728]

### Παρατηρήσεις:

- Το Update 1 και 2 και το Assignment 1 και 2 έχουν περίπου ίδια απόδοση. Ενώ το Initialization 2 έχει καλύτερη απόδοση απο το Initialization 1.

### 12 Μήνες

new representation	Silhouette	time
Cluster.1	0.155711	
Cluster.2	0.11923	
Cluster.3	0.63223	
Cluster.4	0.110741	

Cluster.5	0.209358	
Cluster.6	0.152479	
Cluster.7	0.359282	
Cluster.8	0.281857	
Cluster.9	0.179881	
Cluster.10	0.155567	
Cluster.11	0.242986	
Cluster.12	0.151214	
Alg:221	0.229211	26.7976

Algorithm: I2A2U2 - Silhouette: [0.178455, 0.199529, 0.156746, 0.183971, 0.200562, 0.182489, 0.209683, 0.185232, 0.21429, 0.2286, 0.154412, 0.175862, 0.189153]

Algorithm: I2A2U1 - Silhouette: [0.155711, 0.11923, 0.63223, 0.110741, 0.209358, 0.152479, 0.359282, 0.281857, 0.179881, 0.155567, 0.242986, 0.151214, 0.229211]

Algorithm: I2A1U2 - Silhouette: [0.106209, 0.128078, 0.0986414, 0.156588, 0.237378, 0.211119, 0.465452, 0.205747, 0.124203, 0.180972, 0.245603, 0.193226, 0.196101]

Algorithm: I2A1U1 - Silhouette: [0.143024, 0.127517, 0.133505, 0.14825, 0.17873, 0.208093, 0.181982, 0.276075, 0.58115, 0.202805, 0.204415, 0.270435, 0.221332]

Algorithm: I1A2U2 - Silhouette: [0.546083, 0.198993, 0.223424, 0.125725, 0.138412, 0.17511, 0.125948, 0.121524, 0.120823, 0.131362, 0.149799, 0.109942, 0.180595]

Algorithm: l1A2U1 - Silhouette: [0.150709, 0.114905, 0.0536952, 0.1822, 0.621618, 0.0757697, 0.164718, 0.136782, 0.183284, 0.123856, 0.179427, 0.286609, 0.189464]

Algorithm: l1A1U2 - Silhouette: [0.0907569, 0.505664, 0.069619, 0.185347, 0.215667, 0.41987, 0.169519, 0.12701, 0.150878, 0.199817, 0.11548, 0.0389058, 0.190711]

Algorithm: l1A1U1 - Silhouette: [0.108979, 0.122751, -0.0566879, 0.153776, 0.113848, 0.109838, 0.498647, 0.151545, 0.0500667, 0.150817, 0.154606, 0.173105, 0.144274]

### Παρατηρήσεις:

- Το Update 1 και 2, το Assignment 1 και 2 και το Initialization 1 και 2 έχουν περίπου ίδια απόδοση.

nn representation	Silhouette	time
Cluster.1	0.235794	
Cluster.2	0.245622	
Cluster.3	0.279508	
Cluster.4	0.301974	
Cluster.5	0.189497	
Cluster.6	0.16711	

Cluster.7	0.21669	
Cluster.8	0.145469	
Cluster.9	0.179572	
Cluster.10	0.223705	
Cluster.11	0.154374	
Cluster.12	0.231515	
Alg:222	0.214236	1.6526

Algorithm: I2A2U2 - Silhouette: [0.235794, 0.245622, 0.279508, 0.301974, 0.189497, 0.16711, 0.21669, 0.145469, 0.179572, 0.223705, 0.154374, 0.231515, 0.214236]

Algorithm: I2A2U1 - Silhouette: [0.153382, 0.15339, 0.328193, 0.238838, 0.0812149, 0.161542, 0.201687, 0.191675, 0.168382, 0.219189, 0.131029, 0.196875, 0.18545]

Algorithm: I2A1U2 - Silhouette: [0.132629, 0.199216, 0.195994, 0.157696, 0.191806, 0.215749, 0.179217, 0.0568916, 0.154116, 0.170286, 0.295816, 0.211962, 0.180115]

Algorithm: I2A1U1 - Silhouette: [0.157515, 0.304264, 0.086341, 0.161847, 0.298795, 0.194164, 0.105363, 0.0899889, 0.169573, 0.267728, 0.266044, 0.142589, 0.187017]

Algorithm: I1A2U2 - Silhouette: [0.258143, 0.17455, 0.159803, 0.303432, 0.182822, 0.177292, 0.218212, 0.104454, 0.178297, 0.272163, 0.147586, 0.188765, 0.197127]

Algorithm: I1A2U1 - Silhouette: [0.210572, 0.174457, 0.188842, 0.158037, 0.154551, 0.133208, 0.185558, 0.211567, 0.198198, 0.0150317, 0.148067, 0.215167, 0.166105]

Algorithm: I1A1U2 - Silhouette: [0.235147, 0.202212, 0.136981, 0.293851, 0.275064, 0.142635, 0.211855, 0.200537, 0.134806, 0.18263, 0.171205, 0.124047, 0.192581]

Algorithm: I1A1U1 - Silhouette: [0.148127, 0.111643, 0.178538, 0.200285, 0.352534, 0.221028, 0.247302, 0.0396687, 0.207867, 0.248372, 0.11619, 0.124992, 0.183046]

### **Παρατηρήσεις:**

- Το Update 1 και 2 και το Assignment 1 και 2 έχουν περίπου ίδια απόδοση. Ενώ το Initialization 2 έχει καλύτερη απόδοση απο το Initialization 1.

### **Γενικές Παρατηρήσεις:**

- Διαλέξαμε να παρουσιάσουμε στο readme τις καλύτερες περιπτώσεις clustering. Οι υπόλοιπες μετρήσεις υπάρχουν στο φάκελο results. Παίρνουμε καλά αποτελέσματα στην κατηγοριοποίηση των 4 εποχών καθώς και των 12 μηνών. Η κατηγοριοποίηση των 4 εποχών έχουν ένα μικρο προβάδισμα καθώς είναι πιο εύκολη η κατηγοριοποίηση 4 συστάδων παρά 12. Για τον ίδιο λόγο οι χρόνοι της κατηγοριοποίησης 4 εποχών είναι καλύτεροι.