|  |  |
| --- | --- |
| new_eap_logo.jpg | **ΕΛΛΗΝΙΚΟ ΑΝΟΙΚΤΟ ΠΑΝΕΠΙΣΤΗΜΙΟ** |

**Εργασία: Ομαδικό Προγραμματιστικό Project**

**Ακαδημαϊκό Έτος :** 2021-2022

**Τίτλος Project:** *Πρόβλεψη Καμπάνιας Μarketing με χρήση αλγορίθμων μηχανικής μάθησης*

**ID Project:** 40

Σκοπός της εργασίας είναι η ανάπτυξη σε Python ενός λογισμικού Μηχανικής Μάθησης με γραφικό περιβάλλον χρήσης, για την πρόβλεψη ανταπόκρισης πελατών σε καμπάνια Marketing.

Περιγραφή Προβλήματος

Μία εταιρεία Μarketing πρόκειται να διοργανώσει μία νέα καμπάνια προώθησης για μία αλυσίδα καταστημάτων πώλησης ηλεκτρικών ειδών μέσω διαδικτύου. H καμπάνια θα στοχεύσει άτομα-πελάτες της αλυσίδας βάσει ιστορικών δεδομένων. Τα δεδομένα αυτά αφορούν προηγούμενη διαφημιστική καμπάνια και περιγράφουν δημογραφικά στοιχεία αλλά και στοιχεία χρήσης του ηλεκτρονικού της καταστήματος της αλυσίδας. Επίσης περιέχουν πληροφορία για το εάν τα συγκεκριμένα άτομα ανταποκρίθηκαν (συμμετείχαν) ή όχι, στην παλαιότερη καμπάνια προώθησης.

Τα ιστορικά δεδομένα αυτά είναι αποθηκευμένα στο αρχείο ‘PastCampaignData.xlxs’ και περιέχουν:

* δημογραφικές πληροφορίες (ηλικία, φύλο, περιοχή κατοικίας),
* πληροφορίες χρήσης μέσων επικοινωνίας (τύπος email, χρήση κινητού τηλεφώνου),
* πληροφορίες χρήσης του ηλεκτρονικού καταστήματος (login – χρήση του λογαριασμού στο ηλεκτρονικό κατάστημα τις τελευταίες 4 εβδομάδες/6 μήνες, αγορές τις τελευταίες 4 εβδομάδες/6 μήνες και σύνολο αγορών),
* πληροφορία για το εάν ανταποκρίθηκε στην καμπάνια,

για κάθε συμμετέχοντα.

Ζητείται να γίνει πρόβλεψη για το εάν τα άτομα που αποτελούν στόχο της νέας καμπάνιας, θα ανταποκριθούν σε αυτή. Τα στοιχεία των ατόμων αυτών είναι αποθηκευμένα στο αρχείο ‘ΝewCampaignData.xlxs’ και αναφέρονται ακριβώς στον ίδιο τύπο δεδομένων με τα ιστορικά δεδομένα με μία διαφορά: η πληροφορία ανταπόκρισης στην καμπάνια δεν είναι διαθέσιμη.

Να γίνει η πρόβλεψη ανταπόκρισης (πρόβλεψη συμμετοχής) των πελατών που στοχεύει η νέα καμπάνια μέσω του μοντέλου k-NN. Αυτό σημαίνει ότι θα πρέπει να χρησιμοποιηθεί αυτή δημοφιλής μέθοδος μάθησης μηχανής (machine learning) έτσι ώστε, ο υπολογιστής να «μάθει» (εκπαιδευθεί) πως συμπεριφέρονται οι πελάτες (μέσω των ιστορικών δεδομένων) έτσι ώστε να προβλέψει πως θα συμπεριφερθούν μελλοντικοί πελάτες (νέα καμπάνια). Η πρόβλεψη αφορά μόνο ένα πεδίο, την ανταπόκριση. Η ακρίβεια της πρόβλεψης που δύναται να κάνει ένας αλγόριθμός μηχανικής μάθησης, μετριέται από συγκεκριμένες μετρικές. Οι τιμές αυτών των μετρικών είναι αναπόσπαστο κομμάτι μίας υλοποίησης λογισμικού μηχανικής μάθησης.

Η υλοποίηση του αλγορίθμου σε Python μπορεί να γίνει με ή χωρίς τη χρήση της σχετικής, έτοιμης βιβλιοθήκης.

**Ζητούμενα:**

Να γράψετε κώδικα στη γλώσσα προγραμματισμού Python ο οποίος θα:

1. διαβάζει τα αρχεία ‘PastCampaignData.xlxs’ και ‘ΝewCampaignData.xlxs’.
2. εκπαιδεύει ένα μοντέλο k-NNχρησιμοποιώντας τα ιστορικά δεδομένα (δεδομένα της προηγούμενης καμπάνιας). Θα πρέπει να θέσετε το κατάλληλο πεδίο ως κλάση.
3. Να προβλέψετε την ανταπόκριση των νέων πελατών στην νέα καμπάνια βάσει του μοντέλου k-NN που εκπαιδεύσατε στο προηγούμενο ερώτημα καθώς και την ακρίβεια των προβλέψεων συνολικά (accuracy και class precision) και για κάθε πρόβλεψη.
4. Να δημιουργήσετε κατάλληλο περιβάλλον διεπαφής για τη διαχείριση των πληροφοριών της εφαρμογής από τον χρήστη.

**Παραδοτέα:**

Παραδίδετε ένα αρχείο zip που περιέχει:

* Τον κώδικά σας.
* Οδηγίες χρήσης της εφαρμογής σας.
* Έκθεση (5-10 σελίδων σε μορφή pdf) που θα εξηγεί: τί κάνατε, πώς το κάνατε, πώς μοιράσατε τη δουλειά, ποια ήταν τα αποτελέσματα. Καλείστε να βάλετε screenshots από την εκτέλεση της εφαρμογής σας .
* Επιπλέον, κάθε φοιτητής ξεχωριστά πρέπει να γράψει μια ατομική έκθεση 1-2 σελίδων που θα εξηγεί τί ακριβώς έκανε ο ίδιος, πόσες ώρες δούλεψε και κατάλογο με βιβλιογραφία (τί μελέτησε, από ποιες πηγές, ακόμα κι αν αυτές δεν είναι βιβλία, αλλά ιστοσελίδες ή video).
* Αρχείο pdf με τις διαφάνειες της παρουσίασής σας.
* Προαιρετικά, μπορείτε να φτιάξετε video με την παρουσίασή σας στο youtube και να αναρτήσετε τον κώδικά σας στο github. Σε αυτήν την περίπτωση βάλτε τα links στο κείμενο της έκθεσής σας.