

SOCCKER DATA ANALYSIS

Μέλη του Project (5^ο Εξάμηνο)

Απατσίδης Ιωάννης 2908

Βλάχος Μιχαήλ 2879

Γιαννουλίδης Απόστολος 2906

Τσεκάλας Γεώργιος Δαυίδ 2888

Ασχοληθήκαμε με την ανάπτυξη αυτής της εφαρμογής στα πλαίσια του Σχολείου του κώδικα.

Τι είναι το Soccer Data Analysis;

Είναι μια Desktop εφαρμογή η οποία αφορά χρήστες που ενδιαφέρονται για στατιστικά δεδομένα αγώνων ποδοσφαίρου 24 πρωταθλημάτων του κόσμου και παρέχει ενδεικτικές προβλέψεις. Η εφαρμογή αντλεί δεδομένα από δύο ιστοσελίδες του διαδικτύου (<http://www.footballresults.org/> , <http://clubelo.com/>) χρησιμοποιώντας web scraping .

Τι είναι το Soccer Data Analysis;

Επίσης η εφαρμογή παρέχει προβλέψεις τόσο για το αποτέλεσμα του αγώνα ανάμεσα σε δύο ομάδες του ίδιου πρωταθλήματος αλλά και τις πιθανότητες που έχει μία ομάδα να χάσει , να νικήσει ή να φέρει ισοπαλία στο επόμενο αγώνα ανεξαρτήτως του αντιπάλου.

Ουσιαστικά για κάθε πρωτάθλημα δημιουργείται ένα νευρωνικό δίκτυο και προπονείται κατά επιλογή το οποίο χρησιμοποιείται για την πρόβλεψη του αποτελέσματος .Ενώ για την πρόβλεψη των πιθανοτήτων νίκης , ήττας , ισοπαλίας ανεξαρτήτως αντιπάλου χρησιμοποιείται ένα νευρωνικό για όλα τα πρωταθλήματα το οποίο έχει προπονηθεί εκ των προτέρων.

Επικοινωνία μεταξύ των μελών

Αρχικά πραγματοποιήθηκε συνάντηση μεταξύ των μελών που είχε σκοπό την διαμοίραση του φόρτου εργασίας σε γενικά πλαίσια . Έπειτα υπήρχε επικοινωνία μέσω Διαδικτυακών διαύλων για την επίλυση αποριών αλλά και περαιτέρω οργάνωση. Ανά διαστήματα έγιναν και επιπλέον συναντήσεις όταν αυτό ήταν αναγκαίο για την συνέχιση της ανάπτυξης της εφαρμογής.

Επικοινωνία μεταξύ των μελών

- Συζήτηση για την υλοποίηση της αρχικής ιδέας
- Ανταλλαγή απόψεων για τις τεχνικές που θα χρησιμοποιηθούν στην υλοποίηση των επιμέρους διεργασιών .
- Ανάθεση καθηκόντων και σχεδίαση ευέλικτου χρονοδιαγράμματος

Επικοινωνία μεταξύ των μελών

- Οργάνωση του κώδικα με βάση των δύο υποομάδων που δημιουργήθηκαν
- Χρήση pair programming
- Δυναμικό testing των επιμέρους στοιχείων
 - Νευρωνικών
 - Εφαρμογής

Απαιτήσεις

ΜΗ ΛΕΙΤΟΥΡΓΙΚΕΣ

- Συνεχής πρόσβαση στο Διαδίκτυο
- Python version 3.5
- Python Library Tensorflow
- Java version 1.8
- Εύχρηστη διεπαφή πλοήγησης
- Για την πρόβλεψη απαιτείται να υπάρχει πόρος με το αντίστοιχο νευρωνικό δίκτυο
- Κατά την διάρκεια του loading να αναμένεται το πέρας της διαδικασίας
- Δικαιώματα πρόσβασης στα δεδομένα των σελίδων
- Ύπαρξη αρχείου με τις αντιστοιχίες ονομάτων των ομάδων στις δύο σελίδες

Απαιτήσεις

ΛΕΙΤΟΥΡΓΙΚΕΣ

- Δυνατότητα πρόβλεψης αγώνα από δύο οποιοδήποτε ομάδες εντός του ίδιου πρωταθλήματος
- Εμφάνιση πιθανοτήτων για κάθε ένα από τα τρία αποτελέσματα
- Εμφάνιση πιθανοτήτων για το αποτέλεσμα κάθε ομάδας ανεξάρτητα αντιπάλου
- Εμφάνιση στατιστικών
- Πληρότητα πληροφοριών για όλα τα διαθέσιμα πρωταθλήματα
- Δυνατότητα εξέλιξης του τρόπου υπολογισμού προβλέψεων

Κώδικας

Υλοποίηση Java project (κυρίως εφαρμογή)

Κλάσεις :

- StartingPage (Αρχικό παράθυρο Διεπαφής)
- TeamStats (Παράθυρο με πληροφορίες ομάδας)
- League (Αναπαράσταση Πρωταθλήματος)
- Team (Αναπαράσταση Ομάδας)
- ReadEloNames (Αντιστοίχιση ονομάτων των ομάδων μεταξύ των δύο ιστοσελίδων)
- BuildData (Παραγωγή αρχείων csv με τα δεδομένα που εισαίρονται ως είσοδο στα python αρχεία train, predict , GeneralPredict)
- NNhandler (Σύνδεση με τα αρχεία python)
- Results (scraping αποτελεσμάτων από την ιστοσελίδα)

Κώδικας Python

- GeneralPredict.py (φόρτωση του νευρωνικού και παραγωγή προβλέψεων)
- nn.py (Προπόνηση του νευρωνικού δικτύου που αφορά τα αποτελέσματα των αγώνων)
- Predict.py (φόρτωση του νευρωνικού πρόβλεψης αποτελέσματος μεταξύ δύο ομάδων και παραγωγή προβλέψεων)
- Parse_Leagues.py (προσομοίωση αγώνων για την παραγωγή του κατάλληλου συνόλου δεδομένων για την προπόνηση του νευρωνικού δικτύου πρόβλεψης πιθανότητας του επόμενου αποτελέσματος μια ομάδας ανεξάρτητα από τον αντίπαλο)

Τελικά συμπεράσματα

- Διεκπεραίωση αρχικών στόχων σε ικανοποιητικό χρόνο
- Περαιτέρω κατανόηση λειτουργίας ομαδικών project και συνεργασίας βάσει των αρχών του Software Engineering
- Διεύρυνση γνώσεων μέσω pair-programming και προβλημάτων που προέκυψαν
- Ανάπτυξη των ατομικών δεξιοτήτων στο πεδίο ενασχόλησης του καθενός