**Αναφορά Εργαστηριακής Άσκησης Χειμερινού Εξαμήνου για το μάθημα Ανάκτηση Πληροφορίας**

Στοιχεία Φοιτητών:

Δημήτριος Κωστορρίζος ΑΜ: 1054419 Έτος Σπουδών: 5o

Λάμπρος Παπαδόπουλος ΑΜ: 1054433 Έτος Σπουδών: 5o

Περιγραφή Υλοποίησης

Η υλοποίηση έχει γίνει σε C# 8.0. Το περιβάλλον υλοποίησης που έχει χρησιμοποιηθεί είναι Visual Studio 2019. Χρησιμοποιήθηκαν τα εξής NuGet packages: Elasticsearch.Net, NEST.

Κατεβάσαμε την το αρχείο zip που περιέχει τα αρχεία για την Elasticsearch από το link: <https://www.elastic.co/downloads/elasticsearch>

Τοποθετήσαμε το unwrapped περιεχόμενο του αρχείου στον φάκελο C: και εκτελέσαμε την Elastic search από το CMD. Επιβεβαιώσαμε ότι η Elasticsearch εκτελείται σωστά στην default διεύθυνση localhost:9200.

**Ερώτημα 1**

Αρχικά ορίσαμε ένα default index, σε πεζά(lowercase) γράμματα, για την σύνδεση μας στο Elasticsearch, ώστε να αποφύγουμε τυχόν σφάλματα στην περίπτωση όπου ξεχνούσαμε να δηλώσουμε το index κατά την αποστολή του request. Η σύνδεση μας στο Elasticsearch χρησιμοποιεί την default επιλογή του Single Node Connection Pooling, καθώς δεν χρειαστήκαμε επιπλέον node την υλοποίηση μας.

Δημιουργήσαμε την κλάση Movies για να αναπαραστήσουμε τις εγγραφές του αρχείου movies.csv. Για να ορίσουμε το πεδίο που θα χρησιμοποιείτε ως document Id, χρησιμοποιήσουμε το class attribute:

**[ElasticsearchType(IdProperty = nameof(MoviesId))]**

Έτσι ορίσαμε το πεδίο MoviesId ως το μοναδικό Id για τις εγγραφές τύπου Movie.

Για να ορίσουμε την μετρική ομοιότητας BM25 για το πεδίο Title, χρησιμοποιήσαμε το property attribute:

**[Text(Similarity = «BM25»)]**

Έπειτα, με ένα match query, λάβαμε ταξινομημένα κατά φθίνουσα σειρά της εγγραφές που περιέχουν τον όρο αναζήτησης, με βάση την μετρική Score της Elasticsearch.

**Ερώτημα 2**

**Ερώτημα 3**

Εγκαταστήσαμε το package Microsoft.ML. Δημιουργήσαμε ένα νέο instance της κλάσης MLContext και θέσαμε το την τιμή του seed σε μία σταθερή τιμή, ώστε να διασφαλίσουμε ότι τα αποτελέσματα θα είναι ντετερμινιστικά.

**Ερώτημα 4**