



ΠΜΣ ΣΤΗ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗ ΤΕΧΝΟΛΟΓΙΑ
Τμήμα Λογιστικής και Χρηματοοικονομικής
Πανεπιστήμιο Μακεδονίας

ΕΝΤΥΠΟ ΥΠΟΒΟΛΗΣ ΠΡΟΤΑΣΗΣ **ΘΕΜΑΤΟΣ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ**

(για ένταξη στην Τράπεζα Θεμάτων Διπλωματικών Εργασιών του Π.Μ.Σ.)

1. ΣΤΟΙΧΕΙΑ ΠΡΟΤΑΣΗΣ

| Πεδίο | Στοιχεία |
|---|---------------------|
| Κωδικός Θέματος (συμπληρώνεται από τη Γραμματεία μετά την έγκριση του θέματος από τη Συντονιστική Επιτροπή) | |
| Ημερομηνία Υποβολής | |
| Προτείνων | ΣΥΜΕΩΝ ΠΑΠΑΔΟΠΟΥΛΟΣ |
| Φορέας Προέλευσης Θέματος (FAC, FTSAI, RES, IND, STU, EXT) ¹ | FAC |
| Κύρια Θεματική Περιοχή (FINTECH, , RISK, AI-DATA,, DLT, REG, GOV, PROG, IND) | AI-DATA |
| Δευτερεύουσα Θεματική Περιοχή (προαιρετικά) | FINTECH |
| Τριτεύουσα Θεματική Περιοχή (προαιρετικά) | |

¹FAC: Μέλος Δ.Ε.Π. ή Διδάσκων του Π.Μ.Σ., FTSAI: Financial Technology and Strategic Artificial Intelligence Laboratory, RES: Άλλη ερευνητική δομή ή ερευνητικό έργο, IND: Επιχείρηση ή οργανισμός, STU: Πρόταση φοιτητή, EXT: Εξωτερικός συνεργάτης ή φορέας.

2. ΤΙΤΛΟΣ ΘΕΜΑΤΟΣ

Τίτλος στα Ελληνικά

Αυτόματη παραγωγή Equity Research Reports

Title in English

Automatic Generation of Equity Research Reports

3. ΚΑΤΗΓΟΡΙΑ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

- Ερευνητική Διπλωματική Εργασία
- Εφαρμοσμένη Διπλωματική Εργασία
- Τεχνολογική Διπλωματική Εργασία
- Διπλωματική σε Συνεργασία με Οργανισμό ή Επιχείρηση
- Διπλωματική Ενταγμένη σε Ερευνητική Δράση

Εφόσον επιλεγεί η τελευταία κατηγορία – Ονομα Ερευνητικής Δράσης

4. ΜΕΘΟΔΟΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ

- Βιβλιογραφική ή Θεωρητική Μελέτη
- Συστηματική Βιβλιογραφική Ανασκόπηση
- Εμπειρική ή Ποσοτική Ανάλυση
- Μελέτη Περίπτωσης
- Συγκριτική Ανάλυση
- Ανάπτυξη ή Αξιολόγηση Τεχνολογικού Συστήματος
- Σχεδιασμός Πλαισίου, Μεθοδολογίας ή Μοντέλου
- Μικτή Προσέγγιση

5. ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ ΤΟΥ ΘΕΜΑΤΟΣ

(ενδεικτική έκταση: 100 έως 250 λέξεις)

Η διπλωματική εργασία αποσκοπεί στη σχεδίαση και ανάπτυξη ενός συστήματος τεχνητής νοημοσύνης που παράγει αυτόματα Equity Research Reports για εισηγμένες εταιρείες, αξιοποιώντας Μεγάλα Γλωσσικά Μοντέλα (Large Language Models - LLMs) σε συνδυασμό με δομημένα και μη δομημένα χρηματοοικονομικά δεδομένα.

Το σύστημα θα συλλέγει οικονομικές καταστάσεις, χρηματιστηριακά δεδομένα, βασικούς χρηματοοικονομικούς δείκτες, ειδησεογραφία και εταιρικές ανακοινώσεις και θα δημιουργεί αναφορές αντίστοιχες με εκείνες που συντάσσουν επαγγελματίες equity analysts. Οι αναφορές θα περιλαμβάνουν επισκόπηση της εταιρείας, ανάλυση οικονομικών αποτελεσμάτων, αξιολόγηση βασικών δεικτών, ανάλυση κινδύνων, εκτίμηση προοπτικών και τεκμηριωμένη επενδυτική σύσταση.

Ιδιαίτερη έμφαση θα δοθεί στη χρήση τεχνικών Retrieval-Augmented Generation (RAG), στην αξιοπιστία των παραγόμενων απαντήσεων, στη μείωση των φαινομένων hallucination και στην αξιολόγηση της ποιότητας των παραγόμενων αναφορών μέσω αντικειμενικών μετρικών αλλά και αξιολόγησης από ειδικούς. Το αποτέλεσμα θα είναι ένα λειτουργικό πρωτότυπο (prototype) που θα μπορεί να υποστηρίξει αναλυτές, επενδυτικές εταιρείες και οργανισμούς χρηματοοικονομικών υπηρεσιών. Η διπλωματική εργασία αποσκοπεί στη σχεδίαση και ανάπτυξη ενός συστήματος τεχνητής νοημοσύνης που παράγει αυτόματα Equity Research Reports για εισηγμένες εταιρείες, αξιοποιώντας Μεγάλα Γλωσσικά Μοντέλα (Large Language Models - LLMs) σε συνδυασμό με δομημένα και μη δομημένα χρηματοοικονομικά δεδομένα.

Το σύστημα θα συλλέγει οικονομικές καταστάσεις, χρηματιστηριακά δεδομένα, βασικούς χρηματοοικονομικούς δείκτες, ειδησεογραφία και εταιρικές ανακοινώσεις και θα δημιουργεί αναφορές αντίστοιχες με εκείνες που συντάσσουν επαγγελματίες equity analysts. Οι αναφορές θα περιλαμβάνουν επισκόπηση της εταιρείας, ανάλυση οικονομικών αποτελεσμάτων, αξιολόγηση βασικών δεικτών, ανάλυση κινδύνων, εκτίμηση προοπτικών και τεκμηριωμένη επενδυτική σύσταση.

Ιδιαίτερη έμφαση θα δοθεί στη χρήση τεχνικών Retrieval-Augmented Generation (RAG), στην αξιοπιστία των παραγόμενων απαντήσεων, στη μείωση των φαινομένων hallucination και στην αξιολόγηση της ποιότητας των παραγόμενων αναφορών μέσω αντικειμενικών μετρικών αλλά και αξιολόγησης από ειδικούς. Το αποτέλεσμα θα είναι ένα λειτουργικό πρωτότυπο (prototype) που θα μπορεί να υποστηρίξει αναλυτές, επενδυτικές εταιρείες και οργανισμούς χρηματοοικονομικών υπηρεσιών.

6. ΣΤΟΧΟΙ ΚΑΙ ΕΡΕΥΝΗΤΙΚΑ ΕΡΩΤΗΜΑΤΑ

Στόχοι

- Μελέτη της βιβλιογραφίας σχετικά με τη χρήση LLMs στη χρηματοοικονομική ανάλυση.

- Σχεδιασμός αρχιτεκτονικής αυτόματης παραγωγής Equity Research Reports.
- Ανάπτυξη μηχανισμού συλλογής και προεπεξεργασίας χρηματοοικονομικών δεδομένων.
- Ενσωμάτωση τεχνικών Retrieval-Augmented Generation (RAG).
- Ανάπτυξη προτύπου (template) παραγωγής αναφορών.
- Αξιολόγηση της ποιότητας, αξιοπιστίας και χρησιμότητας των παραγόμενων reports.

Ερευνητικά Ερωτήματα

1. Σε ποιο βαθμό μπορούν τα LLMs να παράγουν αναφορές αντίστοιχης ποιότητας με αυτές των επαγγελματιών equity analysts;
2. Ποια δεδομένα συνεισφέρουν περισσότερο στη βελτίωση της ποιότητας των παραγόμενων αναφορών;
3. Πόσο μειώνεται το ποσοστό λανθασμένων ή μη τεκμηριωμένων πληροφοριών μέσω της χρήσης RAG;
4. Ποιες τεχνικές prompt engineering οδηγούν σε πιο συνεπείς και επαγγελματικές αναφορές;
5. Ποιες είναι οι κυριότερες προκλήσεις και περιορισμοί της χρήσης παραγωγικής τεχνητής νοημοσύνης στον χώρο του Equity Research;

7. ΠΡΟΤΕΙΝΟΜΕΝΗ ΜΕΘΟΔΟΛΟΓΙΑ

Η εργασία θα ακολουθήσει τα παρακάτω στάδια:

1. Βιβλιογραφική ανασκόπηση σχετικά με LLMs, Financial NLP και AI in Finance.
2. Μελέτη της δομής επαγγελματικών Equity Research Reports.
3. Συλλογή οικονομικών δεδομένων από δημόσιες πηγές.
4. Σχεδίαση data pipeline.
5. Ανάπτυξη συστήματος RAG.
6. Σύνδεση του RAG με LLM (π.χ. GPT ή άλλο ανοικτού κώδικα μοντέλο).
7. Σχεδίαση prompts για παραγωγή:
 - Company Overview
 - Financial Analysis
 - Ratio Analysis
8. Υλοποίηση εφαρμογής σε Python.
9. Συζήτηση αποτελεσμάτων και προτάσεις βελτίωση

8. ΔΕΔΟΜΕΝΑ ΚΑΙ ΠΗΓΕΣ ΔΕΔΟΜΕΝΩΝ

Τα δεδομένα θα περιλαμβάνουν:

- οικονομικές καταστάσεις
- ισολογισμούς
- αποτελέσματα χρήσης
- ταμειακές ροές
- χρηματιστηριακές τιμές
- χρηματοοικονομικούς δείκτες
- ειδησεογραφία
- εταιρικές ανακοινώσεις

9. ΠΡΟΑΠΑΙΤΟΥΜΕΝΕΣ ΓΝΩΣΕΙΣ Ή ΔΕΞΙΟΤΗΤΕΣ

- Python
- Βασικές γνώσεις Machine Learning
- Large Language Models
- Prompt Engineering
- APIs
- Χρηματοοικονομική Ανάλυση
- Financial Statements
- Επενδύσεις και Αγορές Κεφαλαίου
- Αγγλική τεχνική βιβλιογραφία

10. ΑΝΑΜΕΝΟΜΕΝΑ ΠΑΡΑΔΟΤΕΑ

- Αναλυτική βιβλιογραφική επισκόπηση.
- Σχεδιασμός αρχιτεκτονικής συστήματος.
- Υλοποίηση λογισμικού παραγωγής Equity Research Reports.
- Σύστημα συλλογής και ενημέρωσης δεδομένων.
- Prompt library.
- Prototype εφαρμογής.
- Αξιολόγηση αποτελεσμάτων.

11. ΕΚΤΙΜΗΣΗ ΔΥΣΚΟΛΙΑΣ

Κλίμακα 1 (χαμηλή) έως 5 (υψηλή)

| Κατηγορία | 1 | 2 | 3 | 4 | 5 |
|-----------------------------|--------------------------|--------------------------|-------------------------------------|-------------------------------------|--------------------------|
| Θεωρητική Δυσκολία | <input type="checkbox"/> | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Προγραμματιστική Δυσκολία | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| Δυσκολία Συλλογής Δεδομένων | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> |

12. ΕΚΤΙΜΗΣΗ ΔΥΣΚΟΛΙΑΣ

- Χαμηλή
- Μέτρια
- Υψηλή

Σύντομη αιτιολόγηση

Η εργασία απαιτεί συνδυασμό γνώσεων χρηματοοικονομικής ανάλυσης, τεχνητής νοημοσύνης και ανάπτυξης λογισμικού. Η κύρια πρόκληση αφορά την ολοκλήρωση πολλαπλών πηγών δεδομένων, τη χρήση σύγχρονων τεχνικών Retrieval-Augmented Generation, την αποφυγή ανακριβών ή μη τεκμηριωμένων πληροφοριών από τα γλωσσικά μοντέλα και την αντικειμενική αξιολόγηση της ποιότητας των παραγόμενων αναφορών. Παράλληλα, απαιτείται ανάπτυξη ενός ολοκληρωμένου πρωτοτύπου που να παράγει συνεκτικά, τεκμηριωμένα και επαγγελματικού επιπέδου Equity Research Reports.