



# Project Proposal

## Product Title Classification

Lecturer: Dr. Muhammad Rafi

Course: Information Retrieval (CS-4051)

Semester: Spring 2024

Section: BCS-6A

06.04.2024

---

Muhammad Salar (K21-4619)

Muhammad Hamza (K21-4579)

Sufyan Khan (K21-3954)



## Executive Summary

Our project is centred on improving the precision of Information Retrieval (IR) systems within the e-commerce sector by refining product title classification and quality assessment. We aim to tackle the challenges posed by the unique nature of e-commerce text data, such as its brevity and contextual sparsity, to deliver clear and relevant search outcomes. The project's methodology involves leveraging sophisticated machine learning strategies, including natural language processing and supervised learning, to preprocess text data, extract features, and train models. The scope of our endeavour encompasses the entire development lifecycle, from data preprocessing to model deployment, focusing on enhancing existing research. The design phase is dedicated to constructing a robust model architecture that can be effectively trained on e-commerce datasets. Our development plan is structured in stages, ensuring thorough documentation and alignment with our objectives at each step. Testing will be conducted using both quantitative metrics and qualitative user studies to ensure model efficacy and user satisfaction. This comprehensive approach is expected to yield a refined solution that significantly elevates the performance of IR systems in e-commerce, ensuring accurate categorisation and quality evaluation of product listings.