Inflammation Severity Scoring Using Machine Learning

Muhib Ul Aziz, 4435384, MSc Data Science

***Abstract*— Accurate assessment of inflammation severity from histological samples is essential for diagnosing and monitoring chronic inflammatory diseases. However, manual grading by pathologists is time-consuming and subject to inter- and intra-observer variability. This study develops a machine learning pipeline to predict histological inflammation severity using quantitative inflammatory-cell features extracted from colon tissue samples. The severity score is modelled as a continuous outcome (0–5), with an additional binary inflammation task defined as score <3 versus ≥3. Patient-level cross-validation is employed to prevent data leakage across multiple assessments of the same subject. Experimental results show that models trained on the full feature set achieve strong predictive performance, while a reduced model using at most ten features maintains comparable discrimination with improved interpretability. These findings demonstrate the feasibility of reliable, data-driven inflammation scoring using structured histological features and highlight the potential clinical value of compact and interpretable machine learning models.**

***Index Terms*— Histological inflammation, severity scoring, machine learning, tabular features, feature selection, patient-level cross-validation, binary classification, digital pathology.**

I. Introduction

Histological assessment of inflammation severity plays a critical role in the diagnosis and management of inflammatory diseases such as ulcerative colitis and inflammatory bowel disease. Conventional scoring systems rely on expert visual interpretation of tissue morphology and immune-cell infiltration, introducing subjectivity and variability across observers and clinical centres. Recent advances in digital pathology and machine learning have enabled the extraction of quantitative features from histological slides, offering a more objective and reproducible alternative to manual grading.

This coursework investigates whether machine learning models trained on quantitative inflammatory-cell features can accurately predict histological inflammation severity. The severity score is treated as a continuous outcome (0–5) and is additionally evaluated in a binary clinical setting (<3 vs ≥3). Furthermore, the impact of feature reduction is analysed to determine whether compact models using at most ten features can retain strong predictive performance while improving interpretability and robustness.

The objectives of this study are to:

(i) build a machine-learning model for inflammation severity prediction,
(ii) assess binary inflammation classification performance,
(iii) compare models using full and reduced feature sets, and
(iv) apply patient-level cross-validation to ensure reliable evaluation.

II. Literature Review

Evaluating inflammation and tissue injury on histological slides is a key part of diagnosing and tracking diseases such as ulcerative colitis, inflammatory bowel disease, and steatotic liver disease [3], [5], [7], [8], [9]. Conventional grading frameworks, such as the Geboes et al. ulcerative colitis (UC) scale, classify tissue severity on a 0–5 scale based on features like neutrophil presence, crypt abscesses, epithelial damage, and ulceration, and are commonly used in both clinical research and everyday practice [7]. While these scoring systems have improved inter-observer agreement, substantial variability persists among pathologists, particularly for borderline grades and subtle inflammatory patterns [5], [7]. This challenge has encouraged the adoption of machine learning (ML) and digital pathology as more objective and reproducible alternatives for inflammation assessment [1] – [6], [8], [9].

A. Computational Pathology and Inflammation-Related Histology

Recent studies in computational pathology demonstrate that quantitative features derived from H&E slides can capture immune and stromal patterns that meaningfully predict clinical outcomes [1], [2], [3], [4], [6]. Trahearn et al. analysed over 180 million cells in colorectal cancer cohorts and showed that spatial patterns of lymphocytes, macrophages, and endothelial cells around tumour regions were significantly associated with progression-free survival [1]. Their system used deep learning for cell detection and classification, combined with spatial features such as endothelial-to-cancer cell ratios and immune-cell densities measured within a tumour-associated radius. While their work focused on prognosis rather than severity scoring, it clearly shows that immune and endothelial micro-architecture encodes valuable diagnostic information, supporting the use of inflammatory cell features in ML-based severity models, including those used in this coursework [3], [4], [6], [8].

Foundation models build on this idea by learning general-purpose histology representations from very large and diverse slide datasets [2], [4]. A leading example is PathOrchestra, a ViT-based model pretrained on about 287,000 whole-slide

images from multiple organs, which delivers high accuracy across over 100 downstream tasks, including tumour classification, biomarker prediction, and inflammatory cell detection [2]. These models process images in patches and use attention-based MIL or segmentation networks instead of explicit tabular inputs, aligning with other deep-learning approaches used for inflammation or disease assessment [1], [3], [4], [6], [9]. Although their computational scale is beyond what can be implemented in this coursework, they clearly demonstrate that histology contains enough signal for reliable automated diagnosis, with learned representations of immune infiltration and tissue architecture proving highly informative [1] – [6], [9].

More specialised research has targeted inflammatory bowel disease (IBD) and other inflammation-related disorders [3], [4], [7] – [9]. Furlanello et al. proposed IBD-AI, a system that uses deep learning to detect and count plasma cells in the basal mucosa of colonic biopsies, linking these counts to IBD diagnosis [3]. Because basal plasmacytosis is a well-established marker of IBD, their finding of higher plasma-cell densities in ulcerative colitis and Crohn's disease supports the clinical value of quantitative inflammatory-cell features [3], [7]. Ayad et al. similarly trained attention-based MIL models to classify fetal inflammatory response severity in umbilical cord WSIs across three grades [4]. Their ensemble models achieved high balanced accuracy and AUC, and attention maps consistently highlighted neutrophil-dense arterial walls and Wharton's jelly as important regions, again demonstrating the diagnostic strength of neutrophil-driven inflammation [4], [7].

For liver disease, Abdurrachim et al. assessed an AI-powered digital pathology solution that applies SHG/TPEF imaging with quantitative fibrosis traits to assist MASH fibrosis staging [5]. Although the internal algorithm was not interpretable, AI-derived continuous fibrosis scores significantly improved agreement among pathologists and reduced adjudication needs, showing the practical benefits of quantitative histology-based severity metrics [5].

In addition to fibrosis studies, Nakatsuka et al. used deep learning on non-cancerous steatotic liver biopsies to predict whether patients would later develop hepatocellular carcinoma (HCC) [6]. Their model, trained with patient-level cross-validation and tested on an external dataset, identified subtle but meaningful features, including nuclear atypia, immune infiltration, and fibrosis patterns linked to future cancer risk [6]. Collectively, the literature demonstrates that automated histology can (i) capture inflammatory and structural tissue changes [1]–[6], [9], (ii) improve consistency between pathologists [3], [5], [7], and (iii) deliver useful diagnostic and prognostic information [1], [4], [5], [6]. These findings motivate a comparable ML framework for colon inflammation scoring.

B. Machine Learning for Clinical Severity and Inflammation Scoring

While many histology-based systems rely on deep learning directly on images, several studies have applied classical ML models such as Random Forest (RF), XGBoost, and Logistic Regression to structured clinical or biological features for severity prediction [8], [10],[11]. Chang and Chen used self-reported symptoms, treatment, and contextual features from a chronic-illness monitoring platform to classify patients into three severity levels using Logistic Regression, Support Vector Machines, Random Forest, XGBoost, and LightGBM [10]. Gradient boosting models achieved the highest F1-scores (≈0.85), followed by Random Forest and Logistic Regression, and feature-removal experiments showed that treatment and symptom variables were the most predictive [10]. Although this work is not histology-based, it demonstrates that XGBoost and Random Forest are strong baselines for multi-class severity classification with tabular data [10]–[12], and that model comparison can reveal which algorithms are most appropriate for a given task.

Shim et al. took a mechanistic approach, combining QSP simulation outputs with multinomial Logistic Regression to predict IBD severity scores such as Mayo, MES, and CDAI using simulated biomarker data [8]. Their models achieved good sensitivity and specificity for separating different severity classes, indicating that supervised learning can translate biomarker patterns into recognised clinical scoring systems [8], [10] – [12]. This idea aligns closely with predicting histological severity from inflammatory cell counts, although their input features are simulated inflammatory biomarkers rather than observed histological measures [3], [7], [8].

Other studies in medical machine learning also highlight the strength of XGBoost and Random Forest in predicting disease severity or patient risk [10], [11]. Zou et al. used these models, along with Logistic Regression and others, to stratify diabetes patients based on treatment response and albuminuria progression, finding that XGBoost gave the most accurate HbA1c-reduction forecasts [11]. Salehinejad et al. built an early-warning model for hospitalised patients using XGBoost and longitudinal deterioration-index data, achieving AUCs up to 0.94 and showing good performance across several hospital sites [12]. These studies, although not specifically focused on histology or inflammation, provide robust evidence that ensemble tree methods and Logistic Regression are effective and widely accepted tools for clinical classification and risk modelling [10] – [12].

In the endoscopy domain, Xue *et al.* proposed an automatic severity grading system using confocal laser endomicroscopy (pCLE) images for UC [9]. A convolutional neural network based on EfficientNet-B4 classified images and videos into multiple grades and also performed binary classification between low and high activity, achieving very high AUROC (up to 0.999) and showing strong correlation with histopathology [9]. These findings support results from other imaging and histology research showing that ML-based systems can accurately distinguish levels of inflammatory activity [1], [3] – [6], [9]. While their model uses image features rather than tabular inputs, the overall task is highly comparable

to that of this coursework—multi-class severity grading plus a binary low/high activity classification [8], [10] – [12].

C. Evaluation Methodology and Patient-Level Cross-Validation

A major challenge in clinical machine learning is ensuring that reported performance reflects true generalisation to unseen patients, rather than memorisation of patient-specific patterns. Oner et al. investigated this issue by comparing slide-level and patient-level data splitting when training deep-learning models to predict gene-mutation status from histopathology images [13]. They demonstrated that when slides from the same patient are distributed across both training and test sets, models show artificially inflated performance; enforcing patient-level separation yields more realistic but lower AUC values [13]. Similar principles are followed in other recent multimodal survival and risk-prediction studies, where patient-level cross-validation is used to avoid information leakage [14]. These findings directly support the requirement in this coursework that all rows belonging to the same patient (PatID) must be assigned to the same cross-validation fold [13], [14]. To summarise the reviewed studies and highlight differences in data types, modelling approaches, and performance, Table 1 provides a comparative overview of key related work.

Study	**Data & Task**	**ML Approach**	**Key Contribution**	**Limitation**
Trahearn et al. [1]	Histology (WSI), prognosis	DL + spatial analysis	Demonstrated prognostic value of immune spatial patterns	Not severity scoring
Yan et al. [2]	Histology (WSI), multi-task	ViT foundation model	Strong generalisation across >100 tasks	Computationally intensive
Furlanello et al. [3]	Histology, IBD diagnosis	CNN-based cell detection	Clinically interpretable plasma-cell features	Binary diagnosis only
Ayad et al. [4]	Histology (WSI), severity	Attention-based MIL	Localised neutrophil-driven inflammation	Image-only approach
Abdurrachim et al. [5]	Digital pathology, fibrosis	AI-assisted scoring	Improved inter-pathologist agreement	Limited interpretability
Shim et al. [8]	Simulated biomarkers, severity	Multinomial LR	Maps ML outputs to clinical scores	Simulated inputs
Chang & Chen [10]	Tabular clinical data	LR, RF, XGBoost	Strong tabular ML baselines	No histology data

Table 1: Summary of Related Work on Machine Learning–Based Inflammation and Severity Assessment

D. Identified Gaps and Relevance to the Current Study

Across the reviewed literature, several clear trends emerge. First, histology-based AI systems consistently demonstrate that inflammatory cell densities, spatial organisation, and tissue architecture provide highly informative signals for diagnosis, severity assessment, and prognosis in conditions such as IBD, colorectal cancer, fetal inflammatory response, and liver disease [1] – [6], [9]. Second, classical ML models like Random Forest, XGBoost, and Logistic Regression perform strongly on structured medical data for severity and risk prediction tasks, often outperforming simpler baselines and enabling interpretable feature-importance analyses [8], [10] – [12]. Third, patient-level cross-validation is now recognised as a best practice to prevent data leakage and over-optimistic performance estimates [13], [14].

As shown in Table 1, most existing approaches rely on deep learning applied directly to whole-slide images or specialised imaging modalities, with limited use of structured inflammatory-cell features. However, several gaps remain that motivate the design of the present coursework system. Most histology-focused studies either rely on deep convolutional or transformer-based architectures applied directly to whole-slide images, or depend on specialised imaging modalities such as SHG/TPEF or pCLE, which are not routinely available in everyday clinical workflows [1] – [6], [9]. There is comparatively less work on tabular inflammatory-cell features (densities, ratios, co-localisation metrics) as direct inputs to ML models, especially for jointly addressing: (i) multi-class or continuous severity scoring (0–5) [7], [9], (ii) corresponding binary inflammation classification [4], [9], and (iii) explicit comparison between full-feature models and reduced-feature models using at most ten variables. In addition, while ensemble methods and Logistic Regression have been evaluated separately in many clinical ML contexts [8], [10]–[12], few studies systematically compare Random Forest, XGBoost, and Logistic Regression within a single inflammation-scoring framework under strict patient-level cross-validation, a requirement strongly supported by recent methodological analyses [13], [14].

Most existing works evaluate these algorithms separately rather than in a unified inflammation-scoring framework. This leaves a gap that the present coursework aims to fill.

E. Contribution of the Present Work

To address these gaps, this coursework develops a machine-learning pipeline that uses quantitative inflammatory cell features extracted from colon histology, evaluates both classification and regression formulations of the 0–5 severity score, compares multiple model families (Random Forest, XGBoost, Logistic Regression), and analyses the impact of feature reduction on multi-class/continuous severity prediction and binary inflammation detection, All analyses follow patient-level cross-validation to prevent data leakage.

III. Dataset and Problem Formulation

The dataset consists of 106 subjects enrolled in a clinical study, each identified by a unique patient identifier (PatID). For each subject, up to three histological assessments were performed at different time points (baseline, first trimester, and one year), resulting in multiple rows per patient. Some subjects missed one or more visits, leading to incomplete assessments.

Each assessment includes a histological severity score ranging from 0 to 5, representing the degree of inflammation, along with 141 quantitative features derived from tissue regions of interest. These features capture inflammatory-cell densities, cell-to-cell ratios, spatial co-localisation metrics, and measures of variability across multiple tissue regions.

Two prediction tasks are formulated:

- Severity score prediction, where the target is treated as a continuous variable (0–5) and addressed using regression.
- Binary inflammation classification, derived by thresholding the predicted continuous severity scores at 3 (<3 vs ≥3), in accordance with the clinical definition of active inflammation.

IV. Methodology

A. Data Preprocessing

Rows with missing assessments were excluded, and all numerical features were standardised to zero mean and unit variance where required (e.g. Logistic Regression). For tree-based models, raw feature values were used, as these models are invariant to feature scaling. As missing records were removed, no imputation was necessary.

B. Cross-Validation Strategy

To prevent information leakage arising from multiple assessments of the same subject, patient-level cross-validation was applied. GroupKFold cross-validation was used, with the patient identifier (PatID) defining the grouping variable. This ensured that all assessments from a given subject were assigned to the same fold, providing a realistic estimate of generalisation performance to unseen patients.

C. Machine Learning Model Selection

Several machine learning approaches were initially considered for this task, including linear and ensemble-based methods commonly used for tabular biomedical data. Following preliminary experimentation, multiple model families were evaluated, including Logistic Regression, Random Forest, Gradient Boosting, and XGBoost. Random Forest was selected as the primary regression model for detailed analysis due to its lower mean absolute error and highly stable performance under patient-level cross-validation. Boosting-based models were retained for comparative evaluation to assess whether increased model complexity provided additional performance gains.

Using a single model allowed for a more detailed assessment of feature importance, consistency across folds, and the impact of feature reduction, while keeping the overall pipeline concise. This approach reflects the coursework objective of building and critically assessing a robust machine-learning workflow rather than comparing many models at a surface level.

Binary inflammation classification performance was evaluated by applying a clinical threshold to the regression outputs, rather than training a separate classification model.

D. Feature Reduction

To assess the impact of dimensionality reduction, a reduced-feature model was constructed using at most ten features. Feature selection was performed in a leakage-safe manner within each training fold, based on feature-importance rankings derived from the ensemble model. Feature stability across folds was used as an additional criterion, with preference given to biologically plausible inflammatory-cell density and spatial features.

V. Experimental Design

Two experiments were conducted to evaluate the proposed machine learning pipeline.

Experiment 1 (Full Feature Set):
Models were trained using all 141 histological features. Within each patient-level GroupKFold split, the model was trained on the training folds and evaluated on the held-out test fold. Regression performance for predicting the inflammation severity score (0–5) was assessed using the mean absolute error (MAE). Regression outputs were subsequently thresholded to derive a binary inflammation classification (<3 vs ≥3), and corresponding classification metrics were computed.

Experiment 2 (Reduced Feature Set):
Models were trained using at most ten selected features. The same patient-level GroupKFold procedure was applied, with models retrained from scratch in each fold. Regression performance was evaluated using MAE, followed by assessment of derived binary inflammation classification performance.

For both experiments, performance metrics are reported as mean ± standard deviation across all GroupKFold splits, ensuring robust evaluation and preventing patient-level data leakage.

VI. Results

Severity score regression and binary inflammation classification results obtained using patient-level GroupKFold cross-validation are summarised below.

Feature Set	MAE (mean ± SD)
Full feature set (141 features)	0.715 ± 0.128
Reduced feature set (≤10 features)	0.686 ± 0.128

Table 2. Severity score regression performance (MAE ± SD) evaluated using 5-fold patient-level GroupKFold cross-validation.

Feature reduction resulted in slightly improved regression performance, suggesting that removing redundant features can enhance generalisation.

Feature Set	Accuracy	Sensitivity	Specificity	ROC-AUC
Full feature set (141 features)	0.879 ± 0.031	0.834 ± 0.047	0.951 ± 0.057	0.955 ± 0.039
Reduced feature set (≤10 features)	0.876 ± 0.034	0.833 ± 0.045	0.939 ± 0.051	0.958 ± 0.031

Table 3. Binary inflammation classification performance (<3 vs ≥3) derived from regression outputs, reported as mean ± SD across patient-level GroupKFold splits.

The full-feature model achieved marginally higher accuracy and specificity, while the reduced-feature model maintained comparable discrimination with only a modest reduction in performance.

Explicit mapping to coursework brief:

Table 2 addresses severity score regression performance for the full and reduced feature sets (1.a and 2.a), while Table 3 reports binary inflammation classification metrics including accuracy, sensitivity, specificity, and ROC-AUC (1.b and 2.b), all evaluated using patient-level cross-validation.

VII. Discussion

The results demonstrate that quantitative inflammatory-cell features provide a robust basis for automated inflammation assessment. Strong ROC-AUC values for both experiments indicate effective separation between inflamed and non-inflamed cases, reflecting the rich spatial and cellular information encoded in the extracted features.

Notably, the reduced-feature model achieved slightly lower MAE than the full-feature model, while maintaining very similar binary classification performance. This suggests that some of the original 141 features may be redundant or introduce noise, and that restricting the model to a stable subset of informative features can improve generalisation. These findings support the development of compact and interpretable models that may be more suitable for clinical deployment.

A key methodological strength of this study is the use of patient-level cross-validation, which prevents overly optimistic performance estimates caused by repeated measurements from the same subject. Although alternative models could be explored, the chosen model demonstrated stable performance and strong generalisation under patient-level cross-validation, and therefore, further model comparison was not pursued in this study. Limitations include the relatively small dataset size and reliance on pre-extracted tabular features rather than direct image-based learning. Future work should focus on external validation, incorporation of temporal information across visits, and hybrid approaches combining image-based and tabular features.

VIII. Conclusion

This coursework demonstrates that machine learning models trained on quantitative histological features can accurately predict inflammation severity and reliably detect active inflammation. Models using the full feature set achieve strong performance, while reduced-feature models offer a favourable balance between accuracy and interpretability. These findings highlight the feasibility of objective, data-driven inflammation scoring and emphasise the importance of patient-level validation in clinical machine learning studies. Further validation on independent cohorts is required before clinical adoption.

References

[1] *N. Trahearn et al., "Computational pathology applied to clinical colorectal cancer cohorts identifies immune and endothelial cell spatial patterns predictive of outcome," 2025. [Online]. Available: https://pathsocjournals.onlinelibrary.wiley.com/doi/full/10.1002/path.6378*

[2] *F. Yan et al., "PathOrchestra: a comprehensive foundation model for computational pathology with over 100 diverse clinical-grade tasks," 2025. [Online]. Available: https://www.nature.com/articles/s41746-025-02027-w*

[3] *C. Furlanello et al., "The development of artificial intelligence in the histological diagnosis of inflammatory bowel disease (IBD-AI)," 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1590865824007916*

[4] *M. A. Ayad et al., "Deep learning for fetal inflammatory response diagnosis in the umbilical cord," 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0143400425001225*

[5] *D. Abdurrachim et al., "Utility of AI digital pathology as an aid for pathologists scoring fibrosis in MASH," 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S016882782402734X*

[6] *T. Nakatsuka et al., "Deep learning and digital pathology powers prediction of HCC development in steatotic liver disease," 2025. [Online]. Available: https://journals.lww.com/hep/fulltext/2025/03000/deep_learning_and_digital_pathology_powers.24.aspx*

[7] *K. Geboes et al., "A reproducible grading scale for histological assessment of inflammation in ulcerative colitis," 2000. [Online]. Available: https://gut.bmj.com/content/47/3/404*

[8] *J. V. Shim et al., "Combining mechanistic modeling with machine learning as a strategy to predict inflammatory bowel disease clinical scores," 2025. [Online]. Available: https://www.frontiersin.org/journals/pharmacology/articles/10.3389/fphar.2025.1479666/full*

[9] *P. Xue et al., "An automatic severity grading system using confocal laser endomicroscopy to evaluate inflammatory activity of ulcerative colitis: a prospective study," 2025. [Online]. Available: https://www.nature.com/articles/s41598-025-21968-6*

[10] *Y. Chang and X. Chen, "Estimation of chronic illness severity based on machine learning methods," 2021. [Online]. Available: https://onlinelibrary.wiley.com/doi/epdf/10.1155/2021/1999284*

[11] *X. Zou et al., "The efficacy of canagliflozin in diabetes subgroups stratified by data-driven clustering or a supervised machine learning method," 2022. [Online]. Available: https://link.springer.com/article/10.1007/s00125-022-05748-9*

[12] *H. Salehinejad et al., "Novel machine learning model to improve performance of an early warning system in hospitalized patients: a retrospective multisite cross-validation study," 2023. [Online]. Available: https://www.thelancet.com/journals/eclinm/article/PIIS2589-5370(23)00489-3/fulltext*

[13] *M. U. Oner et al., "Training machine learning models on patient-level data segregation is crucial in practical clinical applications," 2020. [Online]. Available: https://www.medrxiv.org/content/10.1101/2020.04.23.20076406v1*

[14] *Y. Hu et al., "Deep learning-driven survival prediction in pan-cancer studies by integrating multimodal histology-genomic data," 2025. [Online]. Available: https://academic.oup.com/bib/article/26/2/bbaf121/8089949*

Originality & Use of Generative AI Statement

I understand that to use the work and ideas of others, including AI generated output, without full acknowledgement, is academic misconduct. I confirm that this coursework submission is all my own, original work and that all sources, summaries, paraphrases and quotes are fully referenced as required by the LSBU Academic Regulations.

DECLARATION OF AI USE:
I ***DID*** use Generative AI technology in the development, writing, or editing of this assignment.
(delete as appropriate)

If you did use Generative AI, please provide detailed responses to the following items:

1. **Specify the tools (e.g. ChatGPT, Copilot..) and the purposes for using Generative AI technology in this assignment. Clearly explain how the AI technology assisted in the development, writing, or editing processes?**

   I used Generative AI (ChatGPT) as a support tool to clarify my coursework tasks what I need to do and improve my academic tone, grammar and sentence structure. One more thing I used it to check IEEE style, research demo paper so I can write in same style. Seek clarification on standard machine learning concepts and workflows (e.g. cross-validation, evaluation metrics), without generating original analysis or results.

   Generative AI was not used to design the machine learning models, perform data analysis, implement experiments, or interpret results.

2. **Outline the sections or parts of the assignment that were developed, written, or edited with the assistance of Generative AI technology?**

   Generative AI was used only for wording refinement, improve clarity, grammar and sentence structure and also for proof reading. For the reference section I also use this so I can write the reference in IEEE style
   - parts of the literature review (language clarity and structure only)
   - the methodology, results, and discussion sections (grammar and academic phrasing only)
   - assistance in formatting references in IEEE style

   All the remaining work like critical analysis, dataset preprocessing, model selection, coding, experimentation, evaluation, and interpretation of results, finding the different research paper, extracting the main points, and sorting it were completed independently.

3. **Provide 2-3 examples of prompts you used when using AI tools for this assignment?**
   Example prompts included:
   - Hey, can you show me how to write a reference in IEEE style
   - I am confused about the coursework steps can you explain what I need to do next?
   - Kindly check this paragraph and improve grammar and sentence structure while keeping the meaning same.

4. **Reflect on how Generative AI technology contributed to the assignment, including its limitations and advantages in the context of the coursework?**

Generative AI assisted with improving clarity, structure, and readability of written text. It did not replace academic thinking, data analysis, model selection, or critical evaluation it can be done by only a human mind AI just assist you and a key limitation of Generative AI is that its outputs must be carefully verified, as it can sometimes produce incomplete or inaccurate information or fail to identify appropriate academic sources. Therefore, all technical decisions, coding, experimental results, and references were independently verified and produced by myself using authoritative academic sources such as Google Scholar.

By including this statement in my coursework submission, I attest that the information provided in this Originality and use of Generative AI Statement is accurate and complete to the best of my knowledge. I understand that providing false information is a violation of the LSBU Academic Regulations and may result in academic and/or disciplinary consequences.