Guidelines for writing Paper

Prepared By: Md Al-Imran and Md. Ehsanul Haque

You have to follow IMRAD Format when you write your paper.

IMRAD= Introduction, Methods, Results, and Discussion

1. Introduction

In Introduction you should add and answer these
- o Provide some initial data or a compelling hook related to the topic.
- o Clearly state the research problem or question.
- o Explain why it is important to address this problem.
- o Present the context or background of the study.
- o Identify the gaps in existing research that your study aims to fill.
- o Specify your contributions.
- o Outline the research objectives or hypotheses.
- o Organization of the paper

Complete writing this within half page.

EXAMPLE

1. **Data or Hook**

   Include an attention-getting hook to attract the readers. This may be an applicable statistic, a startling fact, or an interesting question concerning your study. In case there are some important data points indicating the importance of the issue you are solving, consider mentioning them.
   Chronic kidney disease (CKD) is a current problem affecting more than 10% of people in the world but it is not yet easy to detect it at an early age. Out of every 100 people with CKD, only half of them are detected before the disease progresses to the severe stage, which is not ideal."

2. **Research Problem or Question**

   It Explicitly state the problem that is being attempted to address by your research. This is the major question on which your study should be focused. Confirm that it goes with the hook you have given and it sets context to the rest of your paper.

   Citing the further technical improvements in the field of medical imaging, early diagnosis of CKD is still a significant issue. The present-day methods of diagnosis are not invasive enough and are not effective at detecting the disease at its earliest stage, posing a need in

the specific timely treatment. What kind of non-invasive diagnostic tool using medical imaging and machine learning can be developed in CKD that is more precise?

3. **Importance of Addressing the Problem (40-50 words)**

   Explain why solving this problem is critical.
   Early diagnosis of CKD can help to eliminate the development of the disease, enhance patient outcomes, and cut costs associated with healthcare. The early diagnosis may be facilitated by non-invasive interventions that will facilitate greater access to healthcare in at-risk populations, which will decrease the long-term costs of healthcare systems."

4. **Context or Background of the Study (60-75 words)**

   Write the background or background information of your research. This should comprise of a brief introduction of current knowledge or research in the field, and important developments or trends. State any applicable theories, technologies, or models on which your work is based.

   Machine learning has brought high-tech changes in many industries and specifically in oncology and cardiology where it provides machine-based diagnostic solutions. The use of it in the detection of CKD, particularly with the aid of the medical imaging process is underdeveloped, though. Though research efforts have been concentrated in the area of blood and urine biomarkers, much could be done to develop more effective early diagnosis through the employment of UA-based image-based artificial intelligence solutions, which are fast and suggestive of more accurate diagnosis.

5. **Gaps in Existing Research (40-50 words)**

   Determine what the current research lacks or what the research gaps are that your research is attempting to fill. This makes your research look like something meaningful to the field.
   Available studies are mostly based on conventional machine learning algorithms with non-imaging data to detect CKD. There are few applications of deep learning methods to medical images, and fewer applications integrate explainable AI, which is essential to clinical acceptance and trust."

6. **Your Contribution (40-50 words)**

   State what your research aims to contribute to the existing body of knowledge. What is unique about your approach or findings? How does it build on or challenge previous work?
   The current study will address the research problem by developing model based on deep learning applicable to kidney images with the objective of diagnosing CKD at an early stage through analysis. It aims to enhance diagnosis performance and model

interpretability by using explainable AI techniques, which can lead to trust in healthcare solutions powered by AI and aided by clinical application."

7. **Research Objectives or Hypotheses (40-50 words)**

   Conclude the introduction by clearly outlining your research objectives or hypotheses. These should be directly tied to the research problem and guide the direction of your study.
   The main goal of the study will be the construction of an AI model to predict early CKD based on the convolutional neural network (CNN) approach to kidney images. We suspect that the implementation of explainable AI techniques will increase diagnostic accuracy and model interpretability, comparing to the older diagnostic methods.

8. **Paper Organization**

   This paper is organized as follows: Section 2 presents the methodology, including data preprocessing, model architecture, and training details. Section 3 outlines the results, followed by Section 4, which discusses the findings and their implications, before concluding with a summary and future research directions in Section 5.

2. Literature Review

What is a Literature Review?
- A summary and analysis of existing research on a specific topic.
- Identifies trends, gaps, and key findings in the field.
- **Define the Scope**
- Clarify the focus and boundaries of your review.
- **Search for Relevant Sources**
- Use academic databases for peer-reviewed journals, books, and conference papers.
- **Organize the Review**
- Chronologically, thematically, or methodologically.
- **Summarize Key Studies**
- Focus on objectives, methods, findings, and conclusions.
- **Critically Analyze**
- Compare studies, highlight strengths, weaknesses, and contradictions.
- **Synthesize Findings**
- Link ideas across studies, showing connections and gaps.
- **Cite Sources Properly**
- Follow the appropriate citation style (APA, IEEE, etc.).
- **Be Concise**
- Focus on relevant studies and avoid unnecessary details.
- **Conclude with Research Gaps**
- Identify gaps your research will address and suggest future directions.

For example, when you write a review for a paper you have to follow this approach—Author et al. /In this study/ In this work -----------------→ What method they used in their paper-----→ what's their best or proposed model accuracy ---------------→ find a potential gaps and write it.

The review of each paper length will be 90-100 words. Add 8-10 paper review if it is a conference paper, you can also add more concisely.

3.   Methodology

- **How do you collect the data? Discuss the columns and rows, etc.**
- **Are there any duplicate values? If so, how do you handle them?**
- **How do you encode categorical columns?**
- **Are there any missing values? If so, how do you handle them?**
- **Are there any outliers? How do you check for and handle them?**
- **How do you select features or determine feature significance?**
- **How do you split the data into training and testing sets?**
- **Is there data imbalance? If so, how do you handle it?**
- **Which model(s) do you train?**
- **Is there any model you propose like stacking or voting? If so discuss how it works pseudocode or etc.**
- **What parameters do you use?**
- **Do you use any explainable AI (XAI) techniques or develop any web applications?**
- **Which evaluation metrics do you use to asses the performance of model?**

EXAMPLE

o **Data Collection & Description:**
"The dataset is collected from [source], consisting of columns such as age, gender, medical history, and test results. The rows represent individual patient records."

o **Handling Duplicates:**
"I check for duplicate values using df.duplicated() and remove them using df.drop_duplicates() to ensure clean data."

o **Encoding Categorical Columns:**
"Categorical columns are encoded using LabelEncoder for ordinal data or OneHotEncoder for nominal data to convert them into numerical format."

o **Handling Missing Values:**
"Missing values are imputed using the mean for numerical columns and the mode for categorical columns, or removed using dropna() if necessary."

o **Handling Outliers:**
"Outliers are identified using box plots and Z-scores. They are handled by either capping the values or removing them based on a threshold."

o **Feature Selection & Significance:**
"Features are selected using techniques like RandomForestClassifier or LASSO for feature importance, or by performing correlation analysis to remove redundant features."

o **Splitting Data:**
"The data is split into training and testing sets using train_test_split with a 70-30 or 80-20 ratio."

o **Handling Data Imbalance:**
"I handle data imbalance using oversampling techniques like SMOTE or under sampling the majority class."

o **Model Training:**
"I train models such as Random Forest, SVM, or Neural Networks, depending on the nature of the dataset and problem."

o **Is there any model you propose, like stacking or voting? If so, discuss how it works (pseudocode or etc.)**
"I propose a **stacking model**, which combines multiple base models to improve prediction accuracy. The outputs of these base models are used as input features for a final meta-learner model, which makes the final prediction. This helps leverage the strengths of different models and reduce overfitting."

o **Hyperparameters Used:**
"For Random Forest, I use parameters like n_estimators=100, max_depth=10, and for Neural Networks, learning_rate=0.001, batch_size=32."

o **XAI and Web App:**
"I use SHAP and LIME for model interpretability, and I developed a web app using Flask for real-time predictions and visualizations."

o **Which evaluation metrics do you use to assess the performance of the model?**
"I use multiple evaluation metrics such as **accuracy**, **precision**, **recall**, **F1-score**, and **AUC-ROC** to assess model performance. For multiclass problems, I also use the **confusion matrix** and **macro/micro average F1-scores**."

4. Results and Discussion

1. Structure Overview

- Results: Present the findings of your research clearly and originally. Present the findings of your research clearly and originally. The Results section should contain no interpretation and should be a compilation of facts, numbers, and outcomes. Like training and testing accuracy table, Learning Curves.

- Discussion: Interpret and analyse the results. The section enables you to relate your findings to the research question and the present literature, explain the implications of these findings, and any limitations.

The Results section should be concise, organized, and focused on providing the facts without interpretation. This is the place where you report what you found.

Key Elements to Include in the Results:

1. **Start with a Summary:**

   o Provide a brief overview of the most important findings.

   o For example, "The model achieved an accuracy of 90% on the test set, outperforming the baseline by 15%."

   o

2. **Organize Results Logically:**

   o **Present the results in a structured way. You can organize them based on different aspects, such as:**

      - Model performance (e.g., accuracy, precision, recall)

      - Cross-validation results (if applicable)

      - Statistical tests (e.g., p-values, AUC)

      - Computational cost (e.g., training time, memory usage)

3. **Use Tables, Figures, and Graphs:**

   o Tables: Use tables for structured data (e.g., confusion matrix, model parameters, performance metrics).

   o Figures/Graphs: Use visual aids like ROC curves, precision-recall curves, or bar plots to represent performance metrics visually.

   o Make sure to label them properly and provide descriptive captions.

4. **Present Performance Metrics:**

- **Include relevant performance metrics such as:**

    - Accuracy, Precision, Recall, F1-Score for classification tasks.

    - AUC-ROC, Confusion Matrix.

    - Loss values, Training Time, Inference Time.

    - If applicable, show cross-validation results and their mean/standard deviation.

5. **Be Clear and Objective:**

    - Avoid giving your interpretations in this section. Focus on presenting the raw data.

Example:

The model has an accuracy of 85.2 percent on the test set, target precision of 0.88, recall of 0.82, and F1-score of 0.85. The accuracy using 5 folds cross-validation was an average of 84.9 and a standard deviation of 1.3 percent. The ROC AUC score was 0.91 showing a great sense of discernment betwixt the classes. The model was trained in 45 minutes and inferencing time for each instance was 0.12 seconds."

The Discussion section allows you to interpret the results of your study, compare your findings with existing research, and draw conclusions. It should be structured clearly to guide the reader through your insights and reflections.

1. Start with a Summary of the Key Findings

- What were the main results?

    - Begin by briefly summarizing the key results of your study, especially those that are most relevant to your research question or hypothesis.

    - Example: "The model achieved an accuracy of 91.5% on the test set, surpassing the baseline performance."

2. Interpretation of the Results

- What do the results mean?

    - Explain what the results imply in the context of your study. Discuss how your findings answer the research questions or hypotheses.

    - Example: "The high accuracy and F1-score indicate that the model is not only accurate but also balances precision and recall effectively, which is critical in medical diagnosis."

- **Unexpected findings?**

    - If you encountered unexpected results, briefly discuss why they might have occurred and how they affect your interpretation.

- o Example: "Surprisingly, the recall rate was higher than anticipated, which suggests that the model might be overly cautious in flagging potential kidney disease cases."

3. Explain the Role of Explainable AI (XAI) Techniques

- Why is interpretability important in your model?

  - o If you used XAI methods like SHAP, LIME, or Grad-CAM, explain their role in making your model more understandable and trustworthy.

  - o Discuss how these techniques help make the model's decisions more transparent, especially in areas like healthcare or finance where trust is crucial.

  - o Example: "We incorporated SHAP to explain feature importance, revealing that age and blood test results were the primary drivers of the model's predictions, which aligns with expert medical knowledge."

4. Compare with Previous Work

- **How do your findings compare to other studies?**

  - o Review and compare your results with similar studies in the field. Highlight any differences or similarities and explain why your results might be better or worse.

  - o Example: "Our findings align with the results of Smith et al. (2020), who also reported high accuracy in kidney disease detection using CNNs. However, our approach improves recall by incorporating data augmentation techniques."

- **What gaps does your study address that previous work may not have?**

  - o Identify any gaps in the existing research that your study addresses, especially in terms of methodology, data, or interpretability.

  - o Example: "While previous studies focused on achieving high accuracy, our approach emphasizes the interpretability of model decisions, a critical factor for clinicians."

5. Conclusion and Future work

- Provide a brief conclusion that reflects on the importance of your findings and their potential applications.
  Example: "In conclusion, our deep learning model, combined with XAI techniques, offers a promising approach for early diagnosis of chronic kidney disease, with potential applications in clinical practice."
- Acknowledge the limitations in your methodology, data, or analysis. This could include dataset size, generalizability, or model complexity.
- Example: "One limitation of this study is the small dataset size, which may affect the generalizability of the model to different populations."
- Suggest areas for future research, whether it's improving model performance, expanding the dataset, or exploring other XAI techniques.

- Example: "Future work should focus on increasing the dataset size and diversity to improve the model's generalizability. Additionally, exploring other XAI techniques, such as LIME, could further enhance model transparency."

<mark>Key Part: ABSTRACT</mark>

An abstract should be a brief (usually 150-300 words) and clear summary of your entire research paper. It should succinctly describe the purpose, methods, results, and conclusion of your study. Typically, an abstract is structured as follows:

1. **Purpose/Background (1-2 sentences):**

   o **Provide a brief statement about the research problem or question.**

   o **Introduce the context and importance of the research.**

2. **Methods (1-2 sentences):**

   o **Describe the methodology, such as the type of model or analysis, and the data used.**

3. **Results (1-2 sentences):**

   o **Summarize the main findings or outcomes of the study.**

4. **Conclusion (1 sentence):**

   o **Conclude by stating the implications of your findings and their significance.**

Example Structure of an Abstract:

1. **Purpose/Background:**
   Start with the motivation for your study and the research question. What is the problem you are trying to solve? Why is it important?

   o Example: " Diagnosis of chronic kidney disease (CKD) at an early state is essential to enhancing the outcomes of patients with kidney diseases, and the current practices are either non-invasive or ineffective. The proposed study will endeavor to create a medical imaging-based model of detecting the early stages of CKD through deep learning techniques."

2. **Methods:**
   Briefly describe the methods used to conduct the research. What kind of model did you use, and what were the key steps of your analysis?

   o Example: " We used a convolutional neural network (CNN) trained on kidney images dataset with data augmentation in place to ensure robustness. We found interpretable AI methods such as SHAP to explain the model predictions."

3. **Results:**
   Summarize the key results of your research, focusing on the main findings or outcomes.

   - o Example: " The model self-learned and performed with an accuracy of 91.5% on the test set even outperforming other baseline models. Application of the SHAP values found out that the most influential factors in the prediction were age and blood test results.

4. **Conclusion:**
   End with a brief statement on the importance of your findings or their implications.

   - o Example: "This model demonstrates that deep learning can be effectively used for early CKD detection, with potential applications in clinical practice for improving patient care.**"**

Example of an Abstract

Abstract

Chronic kidney disease (CKD) ranks among the most common causes of illnesses and death worldwide, and the accuracy and efficiency of currently used diagnostic procedures remain doubtful. In this study, an early prediction model of CKD with computerized medical imaging data will be established based on the deep learning method. A convolutional neural network (CNN) was used with pre-trained model on kidney imagery dataset augmented by data augmentation methods. The model attained an accuracy of 91.5% on the test set which is far beyond what the baseline models are. Explainable AI methods, i.e., SHAP, showed that such features as age and values on blood tests had an important impact in the decision-making process of the model. These observations prove that the combination of deep learning models with explainability methods will allow for greater ease of earlier identification of CKD, with a potential practical application of these methods in practice.

Key Tips for Writing an Abstract:

- **Concise: Keep the abstract within the specified word limit (usually 150-250 words).**

- **No References: Do not include references to other studies or papers in the abstract.**

- **Focus on Key Points: Limit details to the most important findings and methodologies.**

- **write the abstract after completing the full paper to ensure it accurately reflects your study.**

When you take any data or info from any source must cite it