Structure of a Data Science Research Paper

Md. Alamgir Hossain

MSc in ICT, IICT, BUET, BSc in CSE, JUST

Research Instructor, Skill Morph

Lecturer, Dept. of CSE, State University of Bangladesh

Research Assistant, Prince Sattam bin Abdulaziz University, Saudi Arabia

Ex-Lecturer, Dept. of CSE, Prime University

Ex-Research Fellow, ICT Division, People's Republic of Bangladesh

Ex-Research Assistant, IPLab, NAIST, Japan

Skill Morph

Outline of a Research Paper

- Title
- Abstract
- Keywords
- 1. Introduction
- 2. Related Work / Literature Review
- 3. Materials and Methods / Methodology
- 4. Results and Discussion
- 5. Conclusion and Future Work
- References
- Appendix (if needed)

Skill Morph

Title

Clear, Specific, and Reflective of the Core Contribution

***(Include keywords such as “A Novel…”, “An Efficient…”, “Explainable…”, “Federated…”, etc. as appropriate)*

Abstract (200–250 words)

- 1–2 sentences: **Problem and Motivation**
- 1–2 sentences: **Methodology**
- 1–2 sentences: **Results and Key Findings**
- 1 sentence: **Conclusion and Significance**

***(Ensure technical clarity and avoid vague claims)*

Skill Morph

Keywords (4–6)

- Include ***method-specific*** (e.g., “Federated Learning”), **domain-specific** (e.g., “Medical Diagnosis”), and ***innovation-specific*** (e.g., “Explainability”, “Adversarial Robustness”) terms.

Skill Morph

1. Introduction

1.1 **Background and Problem Definition**

- ✓ Present the broader context of the field.
- ✓ Define the ***specific problem*** your research addresses.
- ✓ Clarify its technical or practical boundaries.

1.2 **Motivation and Significance**

- ✓ Explain ***why solving this problem is important***.
- ✓ Discuss its ***real-world applications*** or societal/industrial impact.
- ✓ Highlight what is at stake if the problem remains unsolved.

1.3 **Challenges and Limitations of Existing Approaches**

- ✓ Describe the ***technical or conceptual difficulties*** involved in solving the problem.
- ✓ Review ***existing methods/solutions*** briefly.
- ✓ Explain ***why they fall short*** or where they lack performance or generalizability.

Skill Morph

1. Introduction

- 1.4 **Our Contribution and Innovation**
  - ✓ Clearly state ***what your research proposes*** that is new or better.
  - ✓ Present your main contributions as bullet points.
- A table for presenting ***notations and abbreviations*** used in the paper.
- 1.5 **Paper Organization**
  - ✓ Conclude with a short paragraph describing the ***structure of the rest of the paper***.

Skill Morph

2. Related Work / Literature Review

***(Recent and relevant to **problem domain** + **methods used**).*

- Organized either ***chronologically***, ***thematically***, or by ***methodology***.
- For each paper reviewed:
  - ✓ Objective
  - ✓ Methodology/Model
  - ✓ Datasets
  - ✓ Key Results
  - ✓ ***Limitations / Research Gaps***

***(Use Scopus, Web of Science, SCI/SCIE indexed papers (preferably Q1/Q2). Summarize in a comparative table if relevant).*

Skill Morph

3. Materials and Methods / Methodology

3.1 **Overview of Proposed Framework**

- ✓ System Diagram / Pipeline Figure
- ✓ High-level Explanation

3.2 **Dataset and Data Processing**

- ✓ Source, Size, and Description
- ✓ Preprocessing Steps (missing values, scaling, balancing)
- ✓ Feature Engineering or Feature Selection

Skill Morph

3. Materials and Methods / Methodology

3.3 **Proposed Model / Algorithm**

- ✓ Model Architecture (with diagram if DL, transformer, etc.)
- ✓ Mathematical Formulation or Notation
- ✓ Algorithm Steps (pseudo-code or boxed algorithm)

3.4 **Training Strategy**

- ✓ Hyperparameter tuning (Grid, Random, BO)
- ✓ Loss Functions, Optimizers
- ✓ Early stopping, regularization, dropout
- ✓ Cross-validation method (K-fold, LOOCV, etc.)
- ✓ Computational Resources (e.g., Tesla V100, Colab, GPU specs)

Skill Morph

4. Results and Discussion

4.1 **Experimental Setup**

- ✓ Hardware, Software, Environment
- ✓ Evaluation Protocol

4.2 **Performance Metrics**

- ✓ Accuracy, Precision, Recall, F1-score
- ✓ ROC-AUC, PR-AUC
- ✓ Cohen’s Kappa, MCC, FPR, etc.
- ✓ Training/Test Time, Model Size, Memory Usage

Skill Morph

4. Results and Discussion

- 4.3 **Visualizations**
  - ✓ Confusion Matrix
  - ✓ ROC & PR Curves
  - ✓ Training-Validation Curves
  - ✓ t-SNE / PCA Visualizations (if interpretability relevant)
- 4.4 **Statistical Significance Testing**
  - ✓ e.g., t-test, Wilcoxon, ANOVA for metric comparisons

Skill Morph

4. Results and Discussion

- 4.5 **Comparison with Baselines / Prior Studies**
  - ✓ Include a comparison table with multiple models
  - ✓ Highlight improvements over SOTA
- 4.6 **Error Analysis (optional but impactful)**
  - ✓ Analysis of failure cases
  - ✓ Possible improvements

Skill Morph

5. Conclusion and Future Work

- Recap the goal, approach, and findings
- Highlight significance and impact
- Summarize limitations (honest + constructive)
- Mention possible extensions (e.g., scaling, deployment, real-time use)

Skill Morph

References

- Use ***Zotero*** or ***Mendeley*** for managing
- Follow target journal/conference style (e.g., APA, IEEE, Elsevier)

Skill Morph

Appendix (if needed)

- Additional figures, tables, or pseudocode
- Supplementary datasets / external links
- Ethics Approval Statement (for sensitive data)

Skill Morph

Thank You