A REPORT

ON

AI-driven discipline-specific academic language support for research writing

BY

Prakalp Choubey
2021A8PS3210H

Under the supervision of

Prof. Santosh Kumar Mahapatra

Prof. Aruna Malapati

Submitted in partial fulfillment of THE REQUIREMENTS OF
EEE F376: DESIGN ORIENTED PROJECT

BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI

HYDERABAD CAMPUS

December 2023

ACKNOWLEDGEMENT

I would like to take this opportunity to express my appreciation to the people who have helped me throughout the duration of this project.

I would like to express my gratitude to Prof. Santosh Mahapatra and Prof. Aruna Malapati, who have been very approachable, and have invested considerable effort in ensuring the progress in this project.

I would also like to express my gratitude to my fellow project-mates, who have been extremely co-operative, and have helped ensure that our current research into multi-disciplinary papers goes smoothly.

ABSTRACT

Our project seeks to leverage AI-driven methodologies to fashion a tool that supports researchers in the creation of academic literature. Academic writing, distinguished by its specialized vocabulary and unique style, presents distinct challenges compared to other writing forms. Additionally, this writing style exhibits variations not only across disciplines but also from one journal to another. Consequently, our undertaking aims to develop a tool offering researchers both structural and linguistic guidance. This report delineates our efforts to examine the existence of a shared structural framework in Electrical and Electronics papers through the application of Move Analysis.

Keywords: Natural Language Processing (NLP), Scholarly Writing, Writing Strategies, Sequential Steps

TABLE OF CONTENTS

1. Introduction

Academic writing is a formal and structured mode of communication that conveys scholarly information, ideas, and research findings to a specific audience, typically adhering to established conventions and standards of clarity, evidence-based argumentation, and citation. These established standards may vary as per discipline, as per papers investigated in this project, including Move analysis of research articles across five engineering fields: What they share and what they do not (Maswana et al., 2015) [1].

In the context of writing a research paper, "moves" refer to specific actions or strategies that authors use to effectively communicate their research findings and arguments. These moves help structure and organize the paper, making it clear And comprehensible for readers.

We first aim to verify the differences in the writings of researches of various engineering disciplines through identification of different kinds of moves present in a particular disciplines and also identifying the absence of some moves which help in differentiating a particular journal of an engineering field(say.Electrical Engineering) with a journal of another field (say.Chemical Engineering).Then we plan on to extend our research and develop a platform which helps the research writers to choose the words or sentences based on the discipline in which they are writing the research paper and the kind of particular moves they wish to use as a part of publishing their paper.

2. The Corpus

Our dataset comprises approximately 100 experiment-driven research papers in the field of Electrical Engineering sourced from ten reputable journals. Papers of a survey nature, which primarily explore existing advancements in the field, were excluded from the corpus. These survey-based papers tend to exhibit a somewhat distinct structural pattern compared to those focused on experimental findings. The selection of journals and the corresponding number of papers was determined based on information obtained from the following web pages, which list highly cited articles in the domain of Electrical Engineering:

https://www.scimagojr.com/journalrank.php?area=1700
https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=5

3. Moves: Identifying the move type

There are a total of eleven distinct types of moves, each accompanied by its own set of steps. The comprehensive elucidation of each move, along with its requisite steps, aids in discerning the variances between research journals across diverse Engineering Domains.

Move 1: Presenting the background information

Step 1: Citing established knowledge in the field
Step 2: Addressing the main research problems

Move 2: Reviewing related research

Step 1: Referring to prior research
Step 2: Discussing the limitations of previous research

Move 3: Presenting new research

Step 1: Defining the research purpose
Step 2: Describing the main research procedure and outcomes

Move 4: Identifying source of data and method adopted in collecting them

Step 1: Specifying the source of data
Step 2: Indicating the size of the data
Step 3: Outlining the criteria for data collection
Step 4: Describing the data collection procedure
Step 5: Providing contextual details about the data

Move 5: Describing experimental procedures

Step 1: Identifying the primary research apparatus
Step 2: Detailing the experimental process
Step 3: Specifying the criteria for success

Move 6: Describing data analysis procedures

Step 1: Defining relevant terminologies
Step 2: Outlining the process of data classification
Step 3: Identifying the analytical instrument/procedure
Step 4: Noting any modifications to the instrument/procedure

Move 7: Reporting results

Step 1: Recapitulating data analysis procedures
Step 2: Reiterating research questions
Step 3: Enumerating general findings
Step 4: Enumerating specific findings

Move 8: Commenting on results

Step 1: Interpreting the results
Step 2: Comparing the results with previous studies
Step 3: Evaluating the results (or the research)

Move 9: Emphasizing overall results and their significance

Move b: Explaining specific research outcomes

Step 1: Stating a specific outcome
Step 2: Interpreting the outcome
Step 3: Indicating the significance of the outcome
Step 4: Contrasting present and previous outcomes
Step 5: Pointing out the limitations of the outcomes

Move c: Stating research conclusions

Step 1: Indicating the research implications
Step 2: Proposing avenues for further research

4. Analysis of Various Moves Used In Various Disciplines

Move 1: In the fields of structural engineering, environmental engineering, and chemical engineering, researchers often employed only Move 1, Step 1 before proceeding to the second move (Example 1). Conversely, computer science and electrical engineering articles frequently delved into referencing main research problems, specifically Move 1, Step 2 (Example 2).

Move 2: Writers typically utilized Move 2 in a manner akin to Move 1, with all subdisciplines commonly incorporating "Move 2, Step 1: Reference to previous research." However, this inclination did not always lead to the utilization of "Move 2, Step 2: Reference to limitations of previous research."

Move 3: In all subdisciplines, both steps were conventionally applied.

Move 4: Move 4 concerns the data utilized for experiments, or in cases where there is no experiment, for analysis. This move was employed universally.

Move 5: "Move 5, Step 2: Recounting experimental process" was obligatory in chemical engineering and customary in environmental and electrical engineering. Only in chemical engineering was "Step 4: Stating the hypothesis to test" observed.

Move 6: The variations in the use of Moves 4 and 6, along with certain steps in Move 5, as indicated by examples and questionnaire responses, can be attributed to the diverse types of experiments or studies conducted within a single subdiscipline.

Move 7: With the exception of electrical engineering and computer science, all subdisciplines extensively relied on "Move 7, Step 3." Additionally, "Move 7, Step 4" and "Move 8, Step 1" were notably prevalent in environmental, electrical, and chemical engineering. However, it is worth noting that the frequency of utilization varied significantly from article to article, as indicated by their respective standard deviations.

5. Application Of Moves in Electrical Engineering Research Papers.(An example to identify).

To comprehend the concept of moves and their application, let's examine a tangible research paper and pinpoint the various moves it encompasses. This will afford us a clear understanding of how moves operate.

Title: Design and Optimization of a Solar-Powered Microgrid System for Rural Electrification

Introduction:

Move 1: Establishing the Research Context:

Step 1: Offer context on the cruciality of electrification in rural areas and elucidate the challenges faced.
Step 2: Emphasize the importance of microgrid systems and renewable energy sources in mitigating these challenges.

Move 2: Stating the Research Problem or Question:

Step 3: Clearly articulate the central research question: "How can a solar-powered microgrid system be designed and optimized for effective rural electrification?"
Literature Review:

Move 3: Identifying Relevant Literature:

Step 4: Scrutinize existing literature on microgrid systems, solar energy, and their application in rural electrification projects.
Step 5: Discern pivotal technologies, methodologies, and best practices in the design and optimization of microgrid systems.

Move 4: Identifying Gaps:

Step 6: Highlight gaps in the existing literature concerning the specific design and optimization of solar-powered microgrid systems for rural areas.
Theoretical Framework:

Move 5: Establishing a Theoretical Basis:

Step 7: Introduce pertinent theories or models that will guide the design and optimization process (e.g., load flow analysis, energy management algorithms).
Methodology:

Move 6: Describing the Research Design:

Step 8: Elaborate on the steps taken in designing the microgrid system, encompassing component selection, sizing, and integration.
Step 9: Expound on the simulation tools and software employed for modeling and analysis (e.g., MATLAB/Simulink, HOMER).

Move 7: Justifying Methodological Choices:

Step 10: Rationalize the selection of solar energy as the primary power source, underscoring its sustainability and suitability for rural areas.
Data Presentation and Analysis:

Move 8: Presenting Data:

Step 11: Display simulation results, encompassing power flow diagrams, energy generation profiles, and cost estimations.

Move 9: Interpreting Data:

Step 12: Evaluate the performance of the microgrid under diverse scenarios (e.g., varying loads, weather conditions) and discuss the ramifications.
Discussion:

Move 10: Interpreting Results in the Context of the Literature:

Step 13: Compare the findings with existing studies on microgrid systems and rural electrification, highlighting parallels, distinctions, and contributions.

Move 11: Addressing Implications and Significance:

Step 14: Deliberate on the practical implications of the research for real-world applications, emphasizing its potential impact on rural communities.
Conclusion:

Move 12: Summarizing the Main Points:

Step 15: Provide a succinct recapitulation of the key findings, design considerations, and optimization strategies for solar-powered microgrid systems in rural electrification projects.
Highlighting Contributions:

Step 16: underscore the unique contributions of the research, such as innovative design approaches or insights into optimizing microgrid performance.

Recommendations:

Move 13: Providing Practical Suggestions:

Step 17: Proffer recommendations for implementing similar microgrid projects in other rural areas, taking into account factors like community engagement, maintenance, and scalability.
References:

Citing Sources:

Step 18: Ensure all sources and references are appropriately cited in accordance with the IEEE or the pertinent citation style for electrical engineering.
Abstract:

Move 14: Summarizing the Paper:

Step 19: Furnish a concise synopsis of the research, encompassing the research question, methodology, key findings, and implications.
Keywords:

Identifying Key Concepts:

Step 20: Include pertinent keywords relating to solar microgrids, rural electrification, and optimization to enhance discoverability in academic databases.

6. Potential Flaws In Identifying Moves.

In the analysis of electrical research papers, move analysis serves as a valuable tool for comprehending how authors structure their papers, present arguments, and convey their findings. Nevertheless, like any analytical method, move analysis possesses inherent limitations and potential shortcomings.

Subjectivity in Move Identification: The process of identifying moves is not always clear-cut and can be subjective. Different analysts may discern distinct moves within the same text, potentially introducing inconsistencies in the analysis.

Cultural and Disciplinary Disparities: While move analysis presupposes a universal structure for research papers, conventions can diverge significantly across various cultures and disciplines. What constitutes a "move" in one field may not necessarily apply in another.

Complexity of Multifaceted Moves: Certain moves may serve multiple functions concurrently (e.g., articulating objectives while providing background information), making it intricate to neatly categorize them.

Omitting Implicit Moves: Some moves may not be explicitly articulated but are implied through language usage, rendering them more challenging to pinpoint. This can potentially result in an incomplete grasp of the author's communicative intentions.

Post MidTerm Work

After successfully understanding the definition of moves and their uses in the research papers,we extended our work to include data sets which cover a broader part of the research papers so that there is substantial variation and after training the models we get different kinds of moves and the phrases used for different engineering domains.
We also planned to build a fully functional platform which allows the writers to choose a domain of their choice and the kind of move they wish to use based on which our pre trained model will suggest them a list of words and possibly phrases to enhance their writing experience.

Expanding the Dataset

In the initial phase of our project, we opted for a streamlined approach to expedite the process and obtain preliminary results, providing us with insights into the model's performance. Initially, our focus was solely on scanning the "Abstract" section of research papers to identify the moves present, guided by the definitions outlined earlier. This initial scan served as a quick assessment to gauge the efficacy of our model.

Building on the insights gained from this preliminary phase, we proceeded to a more comprehensive data collection effort. We expanded our dataset to include a thorough examination of over 20 research papers, encompassing not only the abstract but also the introduction, body, and other sections of the papers. This expansion aimed to provide a more holistic view of the language structure and moves employed throughout the entirety of research papers.

By broadening our scope to include the complete research papers, we sought a deeper understanding of the nuanced interplay of moves within different sections. This comprehensive approach allowed us to refine our model based on a diverse range of linguistic patterns and structural elements present in various parts of research papers.

Using Paradigms to Identify Moves

` In the context of our project, paradigms refer to specific approaches or models, such as Few Short Learning, utilized in the training process of our AI model. These paradigms serve as frameworks that guide the learning algorithms in understanding and identifying moves within the context of academic writing.
Initially, the outcomes were less than satisfactory, with an accuracy level hovering around 50%. Recognizing the need for improvement, we expanded our dataset to include the analysis of complete research papers, encompassing not only abstracts but also the introduction, body, and other sections. This expansion significantly boosted the accuracy of our model, providing a more nuanced understanding of language patterns and moves throughout the entirety of research documents.

With improved accuracy, we progressed to the next phase of our project – the development of a fully functional web platform. This platform is designed to assist research writers by suggesting words based on their chosen discipline and the specific moves they wish to incorporate into their work.

The Web Platform

We Following the successful refinement of our AI model, we moved forward with the creation of a web platform, incorporating the aforementioned features to assist research writers. The platform's frontend is developed using HTML/CSS, providing a user-friendly interface that aligns with the diverse needs of researchers.

While the frontend has been established, the backend development is currently in progress. This phase involves the implementation of advanced functionalities, database integration, and the incorporation of AI-driven algorithms to seamlessly deliver tailored suggestions to research writers.

Upon the completion of the backend development, our next step is to initiate a trial phase within our college campus. This involves making the platform accessible to research writers within our academic community, inviting them to explore its features and provide valuable feedback. This user feedback will be instrumental in identifying areas for improvement and ensuring that the platform meets the specific needs and expectations of its users.

The iterative nature of our development process allows us to adapt and make necessary changes based on real-world user experiences. Once we incorporate the feedback from the trial phase, our aim is to refine the platform further, ensuring its efficacy and user satisfaction.

Looking ahead, our broader vision involves making the platform accessible to the general

public. This expansion aims to provide researchers across various institutions and disciplines with a powerful tool for improving their research writing processes. By fostering collaboration and incorporating user feedback at each stage, we are dedicated to creating a web platform that not only meets but exceeds the expectations of the research community.

6.Conclusion

In conclusion, this report offers a comprehensive analysis of the identifiable structure within Electrical and Electronics Engineering papers. Delving into existing literature, the study addresses methodological flaws, providing insights into potential reasons for these limitations.The project has made significant progress, particularly with the expansion of our dataset and the completion of extensive work. This expanded dataset has enriched our understanding and contributes to the robustness of our findings.Moreover, in addition to the analytical aspect, the development of a fully functional web platform is a notable milestone. The web platform, incorporating move analysis insights, not only identifies but also optimizes structural and linguistic elements in Electrical and Electronics Engineering papers. The platform is poised to revolutionize the research writing process, providing valuable assistance to researchers.In summary, the project has made substantial strides, and the combination of expanded datasets, completion of extensive work, and the development of a functional web platform positions it as a promising contributor to the improvement of research paper writing processes.

References

1. Maswana, S., Kanamaru, T., & Tajino, A. (2015). Move analysis of research articles across five engineering fields: What they share and what they do not. Ampersand, 2, 1–11. https://doi.org/10.1016/j.amper.2014.12.002

2. Swales, J. (1981, January 1). Aspects of Article Introductions. Language Studies Unit University of Aston in Birmingham.
3. Ramadhini, T., Wahyuni, I., Ramadhani, N., Kurniawan, E., Gunawan, W., & Muniroh, D. (2021). The Rhetorical Moves of Abstracts Written by the Authors in the Field of Hard Sciences.