CSCE 5300: Introduction to Big Data and Data Science

Project Proposal

Project Idea: Automated Essay Scoring (AES) system

Group 18

1. Neelesh Chandra Alapati
2. Hemanth Addepalli
3. Mythri Popuri
4. Jyothirmai Reddybathuni

Introduction

In our schools and colleges, essay writing has always been a golden yardstick to gauge a student's understanding and depth of knowledge on a subject. It's a canvas where students paint their thoughts, one word at a time. But let's spare a thought for our hardworking teachers, who tirelessly go through stacks of essays, one after another. To address this, the integration of big data and data science technologies presents a promising aspect. The high computational power and analytical capabilities of platforms like Hadoop and Spark can revolutionize the essay grading process, making it faster, more consistent, and more objective. Furthermore, the incorporation of Machine Learning (ML) techniques allows for an effective analysis of essays, considering aspects such as grammatical correctness, coherence, and lexical richness.

This project proposes the development of an Automated Essay Scoring (AES) system, a tool equipped to handle large datasets and provide reliable, real-time essay evaluations. By automating the grading process, the AES system aims to assist educators in delivering timely and constructive feedback, thereby fostering a more dynamic and responsive learning environment.

1. Objective

In our project, we aim to make our teachers' lives a bit easier. We plan to build a reliable system that can quickly and accurately grade student essays using Hadoop and Spark. This system will be smart enough to understand and analyze the essays deeply. We are also focusing on creating a simple and user-friendly interface where teachers and students can easily upload essays and get scores without any hassle. Through this project, we hope to bring a new and efficient approach to essay grading, making a positive change in the field of education technology in our country.

2. Motivation

In our schools and colleges, we find our teachers tirelessly grading papers or essays every day. They often find themselves caught in the whirlpool of grading countless essays, a task that demands both time and meticulous attention. This project is not only to ease the grading process but also a golden opportunity for us, the students, to step into the world of Hadoop. It's our chance to roll up our sleeves and learn the magic that Hadoop has in handling data, all while creating a tool that stands as a pillar of support in our education system.

3. Significance

In today's times, technology is making a big splash in schools and colleges, helping to make old teaching methods even better. Our project is a shining example of this change, aiming to lighten the load of teachers who spend a lot of time grading essays. By bringing in a tool that can do this job automatically, teachers can have more time to engage with students in a lively and focused manner, making classrooms a more fun and learning-filled place. Mainly, this project is a great chance for students to learn about Hadoop, a famous tool used to handle big data, helping them learn new skills and get ready for the job market. In this way, our project promises to be a big step in bringing a fresh and modern touch to education, blending the new tech wonders with our age-old teaching traditions, and making the path of learning smoother and more enjoyable.

4. Features

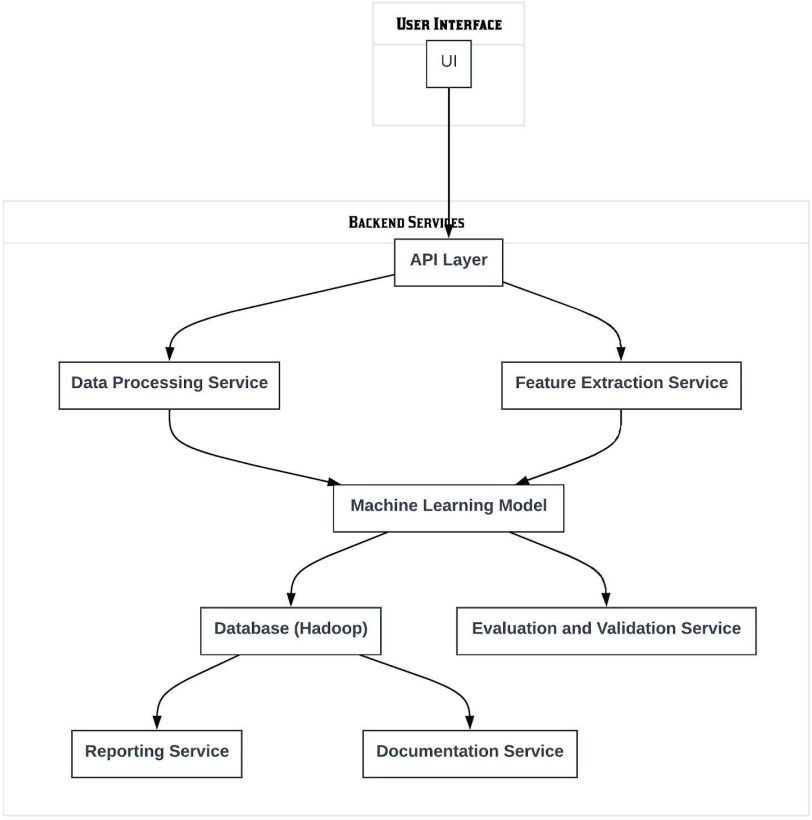
The proposed project is envisioned to be a robust and user-friendly tool with the following key features:

1. Integration with Hadoop: Utilizing the powerful capabilities of Hadoop and Spark to manage and process large datasets efficiently, serving as the backbone of the project.

2. Intelligent Grading System: Working on advanced ML algorithms capable of understanding and analyzing different student essays, thereby ensuring a fair and comprehensive grading process.

3. User Interface: Developing an intuitive interface that facilitates easy navigation and operation, making the process of uploading and retrieving essay scores seamless for the users.

4. Learning Opportunity: Offering students a hands-on experience with Hadoop and Spark, enabling them to learn and apply the principles of this platform in a real-world project setting.



5. Milestones

Our plan is to split into two subgroups. One team will focus on the user interface, while the other will handle data processing. This arrangement will be in place for two weeks before we switch roles. Eventually, we will collaborate to implement ML techniques.

6. Conclusion

In this project, we are focusing on harnessing the power of Hadoop, a well-known framework that excels in processing and managing large data sets. It's like the strong backbone of our project, helping us handle the heavy lifting when it comes to data processing. Alongside Hadoop, we are also planning to integrate Spark, a fast and general engine for big data processing. It works hand in hand with Hadoop, speeding up the data processing tasks and making the system more efficient. To make sure our tool is easy to use, we are putting in efforts to develop a user-friendly interface.

This is a project in which the blend of Hadoop and Spark forms the heart of our project, backed by a simple and intuitive interface with powerful ML techniques, promising a tool that is both powerful and easy to use.