**AIGS 1003**

**MACHINE LEARNING**

**FINAL PROJECT REPORT**

**PROJECT TITLE: SAY MY NAME**

**(PROJECT TEAM: 2)**

**Team Members:**

1. Ninu Basheer
2. Sanjeev Kandel
3. Supriya Polangari
4. Vijay Bhaskar Bonthu
5. Vysakh Mohan Therattil

**ABSTRACT**

The project aims to enhance respect and inclusivity in academic settings by ensuring accurate pronunciation of culturally diverse names. It features an interactive web interface where students submit their names and IDs, and choose phonetic options for correct pronunciation. The backend processes these submissions, offering phonetic suggestions and audio conversions. Unique features include user customization in phonetic selection and a voting mechanism for feedback. This comprehensive approach promotes a more inclusive academic environment by addressing the challenge of name mispronunciation.

**TABLE OF CONTENT**

[1. Introduction 4](#_Toc153356989)

[2. Literature Review 5](#_Toc153356990)

[3. Methodology 6](#_Toc153356991)

[4. Results 9](#_Toc153356992)

[5. Conclusion 12](#_Toc153356993)

[6. Reference Links 13](#_Toc153356994)

# **Introduction**

The project is thoughtfully designed to tackle the challenges associated with accurately pronouncing names originating from diverse cultural backgrounds. Its central objective is not just to recognize every student's name but to ensure that each name is pronounced correctly, ultimately creating an atmosphere characterized by respect and inclusivity.

The journey within this project commences with a seemingly simple yet crucial step. Students engage with an online form, where they provide essential details about themselves. This initial interaction serves as the foundation for capturing the precise information necessary for achieving accurate pronunciation. An option to select their pronouns is also available in this form

Following this, students are presented with a range of phonetic pronunciation options for their names. Multiple suggestions for phonetic representations are presented as radio buttons, offering students a variety of choices. This empowers students to select the pronunciation that best aligns with their preferences and cultural background.

Ultimately, the chosen phonetic representation, whether selected from the provided suggestions is prominently displayed as text in the next page. It is presented alongside the student's preferred name and ID, once this information is securely saved in the database. This ensures that the pronunciation preferences along with the student ID of each student are saved in the database uniquely for future reference.

One noteworthy aspect of this project is its commitment to proper error handling. Whenever an error code is received from the API response, dedicated error handling pages come into play. These pages are strategically incorporated to gracefully manage and communicate errors to the user, ensuring a smooth and user-friendly experience.

In summary, the project's primary goal is to accurately pronounce names from diverse cultural backgrounds, thereby fostering an environment of respect and inclusivity. This is achieved through a structured process that begins with data collection, offers a variety of phonetic options, and ensures proper error handling to enhance the overall user experience.

# **Literature Review**

Numerous academic institutions, including universities and colleges, have conducted extensive research on projects akin to "Say My Name." These initiatives share a common objective: the accurate pronunciation of names, which aligns with our own work. For instance, "The University of Warwick" has initiated a project aimed at investigating whether a lack of familiarity with the pronunciation and spelling of names might pose barriers to effective teaching and learning. (A link to their project is provided below for reference.) Moreover, we've come across several projects with objectives that closely mirror our own, focusing on providing phonetic renditions of names.

One noteworthy example is the project undertaken by "Nottingham Trent University," which shares significant similarities with our initiative. However, our project distinguishes itself by offering a unique feature: customization. We empower users to tailor the phonetic representation of their names according to their preferences. If the system-generated result does not fully meet their expectations, they have the option to choose their preferred phonetics from a set of provided alternatives. This customization aspect sets our project apart and enhances the user experience, ensuring that individuals have control over the pronunciation of their names.

It's worth noting that our project's concept resonates with many online name generator tools. However, a key differentiator is the comprehensiveness of our database, particularly concerning regional names. Many existing tools struggle to provide accurate pronunciation because they lack a diverse range of words and phonetic data related to regional names. In contrast, our project is actively working to expand its database to encompass a wider spectrum of regional names, enhancing our ability to generate precise and culturally sensitive phonetics. This commitment to inclusivity and accuracy sets our project apart in the realm of name pronunciation solutions.

In summary, our project aligns with various academic studies and similar initiatives, all focused on achieving correct name pronunciation. While some projects share common objectives, our distinctive feature is user customization, allowing individuals to select their preferred phonetics. Additionally, our dedication to enriching the database with regional names distinguishes us from many online name generator tools, which often fall short in providing accurate pronunciations due to limited data coverage.

# **Methodology**

**3.1 Frontend**

The frontend of the application has been meticulously designed and styled using Tailwind CSS to provide an engaging and visually appealing user experience. It consists of three distinct user interface pages, each serving a specific purpose and seamlessly interconnected to enhance the user's interaction with the application.

**Input Page**

The initial page serves as the entry point for users. Here, they are prompted to provide two key pieces of information: their preferred name, their pronouns and student ID. The student ID field have validation to check whether maximum limit of 5 digits exceeded or not. Also, it ensures that no alphabetic characters are entered in the field. From the dropdown list user can select a pronoun also. Once the user fills in these details, they can initiate the process by clicking the "Submit" button.

**Data Processing and Phonetics Page**

Upon clicking the "Submit" button on the input page, the frontend triggers a POST request to an API, and it eagerly awaits the API's response. The response received from the API is crucial, and the subsequent actions are determined by the response's HTTP status code. The primary purpose of this page is to present the student's details, including the preferred name and student ID, which were entered on the previous page.

Furthermore, the API response includes phonetic suggestions. These phonetic suggestions are presented in the UI as radio buttons, allowing the user to select the pronunciation that suits them best. The API response encompasses phonetics from various sources, including a G2P (Grapheme-to-Phoneme) library, a standard English dictionary, and, importantly, phonetics from the database. In the case where phonetics for the same preferred name already exists in the database, a sophisticated algorithm selects the top three phonetics based on their associated "Vote" data. After making their selection, the user advances to the next step by clicking a "Submit" button.

**Final Confirmation Page**

The final page in the process serves to confirm the successful submission and storage of the selected phonetics in the database. It displays a success message to reassure the user that their input has been successfully saved. In addition to the confirmation message, the page presents the student's details alongside the saved phonetic pronunciation. To provide an interactive element, an option to play the audio of the saved phonetics is included, allowing the user to hear the pronunciation they've selected. This feature enhances the user's confidence in the accuracy of the recorded pronunciation.

In summary, the frontend of the application offers a well-structured and user-friendly experience, guiding users through the process of inputting their preferred name and student ID, selecting or customizing phonetic pronunciations, and finally confirming the successful storage of their pronunciation choice in the database. The inclusion of audio playback further enriches the user experience by allowing them to audibly validate their selected pronunciation.

**3.2 Backend**

In our system, which leverages the Fast API framework for creating a robust and RESTful web service, the process of data collection commences at the frontend through a user-friendly form. This collected data is subsequently transmitted to a dedicated POST API endpoint, where it undergoes meticulous validation to ensure its compliance with the required structure and format standards.

Once the validation step is successfully completed, the data enters a series of well-defined processing stages within our system:

**G2p Module**

The validated data, which typically includes student names, undergoes transformation within the G2p (Grapheme-to-Phoneme) module. This module effectively converts each name into its respective phonetic syllabic spellings. This transformation is a crucial step to ensure accurate and consistent pronunciation.

**Formatting and Segmentation**

After phonetic conversion, the output undergoes further processing. This includes formatting adjustments to eliminate any numerical elements and organizing the data into a structured list format. Additionally, the data is handed over to the Pippen module, which focuses on the segmentation of words, ensuring that each component is appropriately recognized and processed.

**Audio Conversion with GTTS Module**

The processed data, now in the form of text with accurate phonetic representations, is then directed to the GTTS (Google Text-to-Speech) module. Here, the text data is transformed into an audio format, allowing for the generation of spoken pronunciations.

**Database Storage**

After successfully traversing all these processing stages, the system compiles the information into a well-structured dictionary format. This compiled data is then securely stored within a Postgres database residing on the backend server.

Our database architecture is designed with two core tables:

* **Student Profiles**: This table is dedicated to maintaining comprehensive student profiles. It stores essential student information, including their names, IDs, and other relevant details.
* **Name Selection Logs:** The second table is dedicated to logging the names selected by students on a secondary interface. This logging is of paramount importance for our recommendation algorithm, which evaluates name popularity based on the frequency of selections by users. This data assists in making informed suggestions to users in the future.

The database effectively serves as a repository for both text and audio data. This ensures the availability and accessibility of this data for future retrieval and analysis, contributing to the continued improvement of the system and enhancing its ability to provide accurate and culturally sensitive pronunciations of student names.

**3.3 Challenges Faced**

During project development phase, we encountered several noteworthy challenges related to the generation of phonetic sounds, particularly when dealing with specialized symbols and regional names. While our system demonstrated a high degree of accuracy in generating phonetic representations for English names, it became evident that handling regional names introduced a layer of complexity due to the rich diversity of linguistic patterns.

One of the primary challenges we faced was selecting a suitable phonetic library and devising a method to effectively segment words, especially regional names, into their constituent phonetic components. The diverse array of languages and dialects associated with regional names meant that there was no one-size-fits-all approach. Consequently, developing a system that could adapt to these variations proved to be a formidable task.

Additionally, we encountered discrepancies between the audio output and the corresponding phonetic text representation. While the audio pronunciation often proved to be accurate, the text representation of the phonetics did not consistently align with the audio output. This highlighted the need for enhanced precision and consistency in the phonetic representation process.

These challenges underscore the intricate nature of our project, as it seeks to navigate the complexities of linguistic diversity while striving for accuracy and precision in providing phonetic pronunciations. As we continue to refine our system, we are actively working to address these challenges, ensuring that it can effectively and sensitively handle regional names and deliver accurate phonetic representations consistently.

# **Results**

The overarching goal of this project is to contribute to the creation of more inclusive and welcoming academic environments by addressing the common challenges associated with the pronunciation of culturally diverse names. In academic settings, where individuals from various backgrounds come together, correct name pronunciation is vital for fostering respect and inclusivity.

To achieve this goal, our project provides an intuitive web interface that allows students to easily submit their names and corresponding IDs. What sets our approach apart is the option for students to customize the phonetic pronunciations of their names, tailoring them to their preferences.

Our project architecture is structured to ensure security, error resilience, and a user-friendly experience. The Flask frontend, responsible for user interactions, handles data securely and incorporates robust error handling mechanisms. This ensures that users have a smooth and hassle-free experience while interacting with the system.

On the backend, the project relies on Fast API, a powerful framework, to handle the data processing pipeline. This includes modules for phonetic transformation, audio conversion, and database storage. Phonetic transformation ensures that names are accurately converted into phonetic representations, which are essential for correct pronunciation. Audio conversion allows these representations to be transformed into audio formats for audible playback. The Postgres database serves as the repository for all this valuable data, securely storing user information and their phonetic preferences.

One unique aspect of our system is the incorporation of a 'Voting' mechanism. This feature allows users to provide feedback and vote on the preferred name pronunciations. Over time, this feedback loop contributes to the improvement of recommended name pronunciations, making them more accurate and culturally sensitive.

Our project's web interface is structured into three main HTML pages for an intuitive user experience. The first page is the gateway, where users input their name, pronouns, and student ID, with validation for the ID field. The second page processes this data, presenting phonetic suggestions and offering custom input options. Finally, the third page confirms the user's selections, displaying saved details and playing the chosen phonetic pronunciation. This streamlined, interactive approach ensures user-friendly navigation and effective data handling.

|  |  |
| --- | --- |
| First HTML Page |  |
| Second HTML Page |  |
| Third HTML Page |  |

In summary, our comprehensive approach ensures that names are pronounced accurately and respectfully, thereby promoting diversity and inclusivity in educational spaces. By combining an intuitive user interface, robust data handling, and an adaptive feedback mechanism, our project contributes to a more harmonious and respectful educational environment, where everyone's name is recognized and pronounced correctly.

# **Conclusion**

Our project successfully addresses the challenge of accurately pronouncing names from diverse cultural backgrounds in academic settings, emphasizing respect and inclusivity. It commences with students providing their names and IDs through an online form, a critical step for precise pronunciation. The system then presents phonetic options, allowing students to select the most accurate representation of their name. The final output includes both text and audio formats of the selected phonetic pronunciation, ensuring clarity and inclusiveness.

Technically, the project leverages Flask for front-end development, Python for programming, Fast API and Uvicorn for the server, SQL Alchemy and psycopg2 for database interactions, and GTTS, G2p, Pypen, and Eng-to-ipa for sound processing. The database is managed using Postgres, and Git is utilized for version control.

In conclusion, our project stands as a comprehensive solution that seamlessly integrates front-end and back-end technologies to enhance the pronunciation of names in educational institutions. This endeavour not only bridges communication gaps but also fosters a more inclusive and respectful academic environment. Future developments could include a feature for users to enter phonetics as text, providing a more personalized experience. Additionally, integrating play buttons alongside phonetic suggestions would offer an immediate auditory confirmation, further enriching user interaction and effectiveness. These improvements would significantly augment the tool's usability and precision.

1. **Reference Links**

* <https://warwick.ac.uk/services/dean-of-students-office/community-values-education/saymyname/research/>
* <https://www.ntu.ac.uk/research/groups-and-centres/projects/say-my-name-experiences-and-impacts-of-pronunciation-of-names-in-higher-education-in-contexts-of-culturally-diverse-student-identities>
* <https://us.pg.com/the-name-generator/>
* <https://www.nameshouts.com/>