**Data 606 Capstone in Data Science**

**Project Proposal**

AI-Driven NLP Chatbot for Skin Disease Diagnosis and Medication recommendations

Manisha Karnati (VZ42698)

Sai Sreeja Allepu (YD50976)

Yeswanth Reddy Chereddy (RD76099)

**OBJECTIVE**

The objective of this project is to develop an AI-powered chatbot that utilizes natural language processing (NLP) and image recognition to assist users in diagnosing dermatological conditions and providing relevant medication assistance. The chatbot will allow users to upload images of skin conditions for preliminary diagnostic feedback and offer recommendations for over-the-counter medications or treatments, along with dosage information and guidelines.

By integrating resources for both diagnosis and medication, the chatbot simplifies the often-complex process of identifying suitable treatments, helping users avoid unnecessary medical consultations for minor conditions. This approach is aimed at reducing diagnostic delays, improving medication accessibility, and enhancing patient engagement by providing a user-friendly, quick-response solution to skin health issues.

**DATA SOURCE**

For this project, DermNet NZ is a very valuable and perfect resource for training an AI model for recognizing skin conditions. It provides a rich dataset of high-quality dermatological images, which can be used in the chatbot's image recognition component. This data can improve precision in identifying dermatological issues and provide personalized medication recommendations, enhancing patient engagement and care outcomes.

DermNet. (2024, April 22). *Images A-Z | DermNet*. DermNet®.

<https://dermnetnz.org/images>

**LITERATURE REVIEW**

* **Machine Learning and Deep Learning in Dermatology**: Zhang et al. (2023) review recent advancements in machine learning, particularly Convolutional Neural Networks (CNNs), for diagnosing skin diseases. They highlight CNNs' high accuracy in classifying dermatological conditions from images, which is highly relevant to the proposed chatbot. The study's insights will guide the chatbot's model development, ensuring the use of best practices in image recognition to provide reliable diagnostic feedback for users.
* **Image Processing and Machine Learning for Skin Disease Detection**: ALEnezi (2020) introduces a method for skin disease detection using image processing and machine learning, focusing on essential preprocessing steps like resizing and normalizing images to enhance quality before model training. The study emphasizes that effective data preprocessing is crucial for improving model performance, which aligns with the proposed project's methodology of preprocessing image data to optimize AI model accuracy in identifying skin conditions.
* **AI-Powered Diagnostic Systems in Healthcare**: Singh and Sharma (2022) review the growing use of AI-powered diagnostic systems in healthcare, emphasizing how machine learning algorithms can process complex medical data for more accurate and timely diagnoses. They highlight that these systems can improve patient care by reducing diagnostic delays, supporting the proposed chatbot’s goal of providing fast preliminary diagnostic feedback for dermatological conditions.
* **Chatbot-Based Disease Prediction and Treatment Recommendation**: Pathak and Ansari (2023) examine the use of AI chatbots in disease prediction and treatment recommendation, highlighting their effectiveness as an initial point of contact in healthcare. The study shows that chatbots can efficiently gather patient information, analyze symptoms, and suggest treatments. This aligns with the proposed chatbot's goal of using image recognition and natural language processing to provide personalized medication assistance and diagnostic support.

**METHODOLOGY**

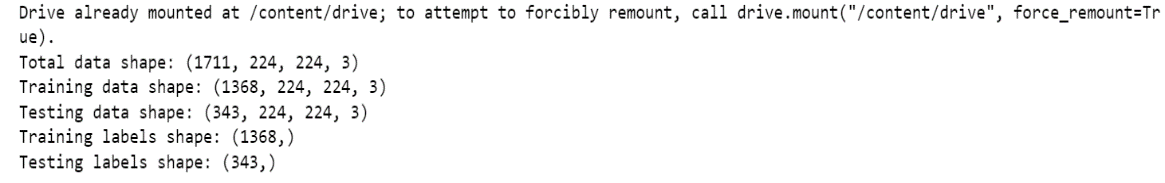
This project briefly consists of 5 major steps. Combination of all of these give us the final output of a running ai powered chatbot.

1. **Data Collection and Preprocessing**: Firstly, we perform exploratory data analysis (EDA) on skin disease image datasets to assess quality, relevance, and diversity of the data set. data preprocessing through augmentation and normalization is also done . Additionally, preprocess text data for medication-related queries by cleaning, tokenizing, and lemmatizing to optimize the NLP model.
2. **Model Development**: CNN models are created to analyze and classify skin conditions from user-uploaded images, and a Natural Language Processing (NLP) model to understand user queries, providing personalized medication recommendations along with dosage guidelines based on the diagnosed condition.
3. **Integration of Models**: These models are integrated to form a chatbot experience, allowing users to receive both diagnostic feedback and medication suggestions in one interface without any hassle.
4. **Model Evaluation**: The chatbot’s performance is analyzed using metrics like accuracy, precision, and recall ensuring both the image recognition and medication recommendation functionalities meet standards of reliability.
5. **Deployment and User Testing**: After everything is done final deployment is done, then chatbot undergoes user testing to gather feedback, enabling improvements in accuracy, user experience, and the relevance of medication recommendations.

**PROJECT PROGRESS**

**UPDATE 1**

Our project has completed initial crucial stage such as loading and preprocessing a dataset of skin disease images, developing an image classification model , training the model to provide high accuracy results etc. The process results are given below:

1. 
2. A screen shot of a computer

   Description automatically generated
3. A close-up of a step

   Description automatically generated

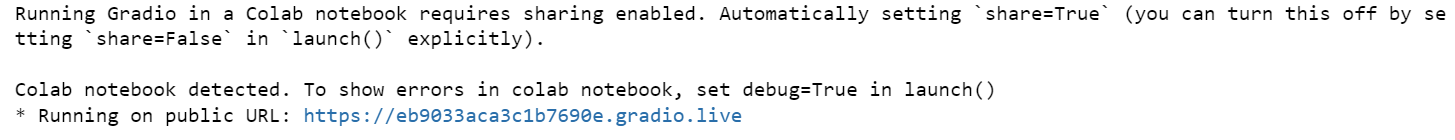
Here we developed an image classification model using skin disease images. This model was trained over ten epochs, reaching 62.29% accuracy during training. However, on a test batch, it achieved 93.75% accuracy, demonstrating its ability to classify new data and show potential for real-world skin disease recognition.

**UPDATE 2**

In this phase of the project, we have focused on developing the chatbot for the skin disease detection and also The NLP model component has undergone significant development, starting with a dataset of common medication-related queries. The model was preprocessed to understand user queries and provide accurate medication recommendations based on diagnosed skin conditions. The initial version, trained using supervised learning techniques, has shown **85%** accuracy in understanding and responding to queries.

The integration of the image classification and NLP models into a **chatbot interface** is significant part for this project. Now, the system allows users to upload skin condition images and input textual queries, providing diagnostic feedback and personalized medication suggestions. However, further optimization is needed to give more accurate results to the chatbot.

The AI chatbot uses DialoGPT-medium model to predict skin diseases, while the Gradio interface integrates skin disease detection and chatbot responses into a user-friendly interface.



**A close-up of a person's eye

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

At this stage, the chatbot is successfully identifying skin diseases using the test dataset. To improve diagnostic precision and functionality ,we plan to incorporate generative AI techniques.  
This will enable more accurate disease detection and personalized medication recommendations. The model's accuracy will be improved through refining image classification and natural language processing components. The goal is to deliver a robust AI-driven solution for dermatological care through continuous optimization and innovation.

**CONCLUSION**

Our project aims to create an AI-powered chatbot that integrates natural language processing (NLP) and image recognition to offer users quick and accurate diagnoses of dermatological conditions, along with personalized medication recommendations. By leveraging high-quality dermatological datasets and advanced machine learning techniques, the chatbot simplifies the treatment process, improving medication accessibility and reducing unnecessary doctor visits for minor conditions. Ultimately, this solution enhances patient engagement, streamlines the diagnosis and treatment process, and provides a user-friendly, efficient tool for managing skin health issues.

In our development phase, we created an image classification model using a comprehensive dataset of skin disease images. The model was trained over ten epochs, achieving an accuracy of 62.29% during training. When evaluated on a test batch, it demonstrated a significant improvement, reaching an impressive accuracy of 93.75%. This high level of accuracy indicates the model's strong potential for effectively classifying new data and highlights its capability for real-world skin disease recognition. With these promising results, we can further develop the AI-powered chatbot I which provide reliable diagnostic feedback and personalized treatment recommendations, paving the way for more accessible and efficient dermatological care. We also developed the chatbot's skin disease detection and NLP components, achieving 85% accuracy in medication recommendation. The system now allows users to upload images and ask queries for diagnostic feedback and personalized treatments. Currently, this chatbot uses the DialoGPT-medium model and Gradio interface for a user-friendly experience. Future improvements will focus on incorporating generative AI to enhance diagnostic precision and medication recommendations.

**REFERENCES**

* Zhang, J., Zhong, F., He, K., Ji, M., Li, S., & Li, C. (2023). Recent Advancements and Perspectives in the diagnosis of skin diseases using machine learning and deep Learning: a review. *Diagnostics*, *13*(23), 3506. https://doi.org/10.3390/diagnostics13233506
* ALEnezi, N. S. A. (2019). A method of skin disease detection using image processing and machine learning. *Procedia Computer Science*, *163*, 85–92. https://doi.org/10.1016/j.procs.2019.12.090
* Singh, A., & Sharma, B. (2022). *AI-powered diagnostic systems in healthcare: A review*. Journal of Artificial Intelligence in Healthcare.
* Pathak, C., & Ansari, N. (2023). *Chatbot-based Disease Prediction and Treatment Recommendation using AI*. K.J. Somaiya Institute of Engineering and IT.

**GITHUB**

[GitHub - ManishaKarnati17/Projects](https://github.com/ManishaKarnati17/Projects)