JAYPEE INSTITUTE OF INFORMATION TECHNOLOGY
SEC- 62, NOIDA

PROJECT REPORT

Course Name: PBL on AI & ML

Course Code: 24B15CS244

Submitted to -
Dr.Deepika Varshney

Submitted by-
Vrinda Gupta - 22101006 (C1)
Khushi Chauhan - 22101011 (C1)
Vaagmi Aggarwal - 22101037 (C2)

INTRODUCTION

"Your skin is the fingerprint of what is going on inside your body."
|Dr. Georgiana Donadio|

From ancient civilizations to modern societies, the desire to achieve flawless, luminous skin has transcended cultural boundaries and persisted across generations. Yet, amidst this timeless aspiration lies a perennial challenge – the daunting task of navigating the labyrinthine landscape of skincare products, treatments, and regimens.

According to the American Academy of Dermatology, acne affects up to 50 million Americans annually, while eczema impacts over 31 million. These statistics underscore the urgent need for personalized skincare solutions that address individual concerns with precision and efficacy.

The skincare industry itself is a vast and ever-expanding ecosystem, with new products, formulations, and trends emerging at an unprecedented pace. From serums and moisturizers to masks and exfoliants,many products have entered the market. The challenge is compounded by the proliferation of misinformation and hype surrounding skincare products, making it increasingly difficult for individuals to make informed decisions about their skincare routines.

Enter DermSelect – a groundbreaking AI-powered platform poised to revolutionize the way individuals engage with skincare products. Born from the confluence of cutting-edge technology and a profound understanding of beauty science, DermSelect stands as a beacon of innovation, offering a solution that is as precise and personalized as it is empowering.

At its core, DermSelect leverages the unparalleled capabilities of AI and machine learning to decode the intricacies of each individual's skin profile with unprecedented accuracy and insight. By seamlessly integrating user-provided images, text data, and advanced algorithms, DermSelect transcends the limitations of traditional skincare analysis, providing recommendations that are tailored to the unique needs, preferences, and aspirations of each user..

As we stand on the threshold of a new chapter in skincare innovation, we invite you to join us in reimagining the possibilities of beauty – one personalized recommendation at a time. With DermSelect as your steadfast ally, the journey towards radiant skin is not merely a destination but a transformative odyssey of self-discovery, empowerment, and celebration. Welcome to DermSelect – where the future of skincare meets the boundless potential of technology, and where beauty is as unique and individual as you are.

REVIEW OF LITERATURE

1) Artificial intelligence in dermatology

Chandler W. Rundle MD , Parker Hollingsworth MD , Robert P. Dellavalle MD, PhD, MSPH

This research paper on "Artificial Intelligence in Dermatology" explores the fundamental concepts of artificial intelligence (AI) and its pivotal role in the field of medicine. The study delves into various methods employed in AI research, elucidating the growth and significance of AI in our rapidly advancing world. Key concepts such as the definition of AI and its application in healthcare are comprehensively discussed. Specifically, the roles of Convolutional Neural Networks (CNN) and Natural Language Processing (NLP) in dermatology are defined, highlighting their contributions to the sector.

Furthermore, the paper emphasizes the critical aspect of data management in AI-driven studies, recognizing the vast volume of data involved and the importance of responsible storage and retrieval practices. This paper serves as a foundational resource for those embarking on projects related to AI in dermatology, providing essential insights into initiating and understanding the application of AI technologies in healthcare.

2) Automatic skin disease diagnosis using deep learning from clinical image and patient information

K. A. Muhaba , K. Dese, T. M. Aga, F. T. Zewdu, G. L. Simegn

This research paper, published on November 25, 2021, focuses on the development of a smartphone-based automated skin disease diagnosis system using artificial intelligence (AI) and deep learning techniques. The study aims to diagnose five specific skin diseases by analyzing clinical images and patient information, including age, gender, anatomical site of the disease, and symptom list. The research involved collecting a comprehensive dataset of clinical images for each skin disease to facilitate model training and evaluation. Patient data, such as age, gender, and symptoms, was integrated into the dataset for analysis. Data preprocessing techniques were applied to customize the diagnosis process based on age-specific disease manifestations. The primary deep learning model utilized in this study was the pre-trained MobileNet-V2 architecture, fine-tuned and trained using the compiled dataset to optimize for accuracy and generalization. Training progress was monitored over multiple epochs, with results analyzed for accuracy and validation loss. By the 45th epoch, the model achieved promising performance with 94.2% accuracy on the training set and 88.3% accuracy on the validation set, accompanied by a minimum validation loss of 0.306. Graphical representations were employed to visualize disease classifications based on age factors, providing insights into disease patterns across demographics. Multi-class classification outputs were tabulated to comprehensively showcase model performance and diagnostic accuracy. The findings highlight the potential of smartphone-based automated skin disease diagnosis using AI and deep learning methodologies, demonstrating high accuracy rates and validating the efficacy of the proposed approach for real-world applications. The integration of patient information alongside clinical images enhances diagnostic precision, enabling tailored recommendations and early interventions. In conclusion, this research presents a novel framework for smartphone-based skin disease diagnosis leveraging AI and deep learning techniques, with future work aimed at further model refinement and potential deployment of the automated diagnosis system in clinical settings.

3) **A Web-Based Skin Disease Diagnosis Using Convolutional Neural Networks**
   Samuel Ayer Bamfo-Sam_Acheampong Addo Philip_Derrick Yeboah1_Nancy Candy Love Nartey_Isaac Kofi Nti

The research paper was published on 19th November, 2018, focuses on building a web based skin disease detection system named ***medilab-plus*** using a convolutional neural network classifier built upon the Tensorflow framework for detecting (atopic dermatitis, acne vulgaris, and scabies) skin diseases. It is estimated that the proposed system will enhance accuracy and offer fasting diagnosis results than the traditional method, which makes this system trustworthy and resourceful for dermatological disease detection. This model classifies three diseases i.e,atopic dermatitis, acne vulgaris, and scabies.The dataset used was preprocessed through data cleaning techniques such as smoothing, aggregation normalization, and attribute construction. Experimental results of the proposed system exhibited classification accuracy of 88% for atopic dermatitis, 85% for acne vulgaris, and 84.7% for scabies. CNN is applied in identifying faces, individuals, street signs, tumors, platypuses, and many other aspects of visual data. The first convolutional layer (CL) then moves to the max pooling layer (PL) stage second convolutional stage until the fully connected neural network is obtained. The combination of the CL and PL forms the ith CNN.

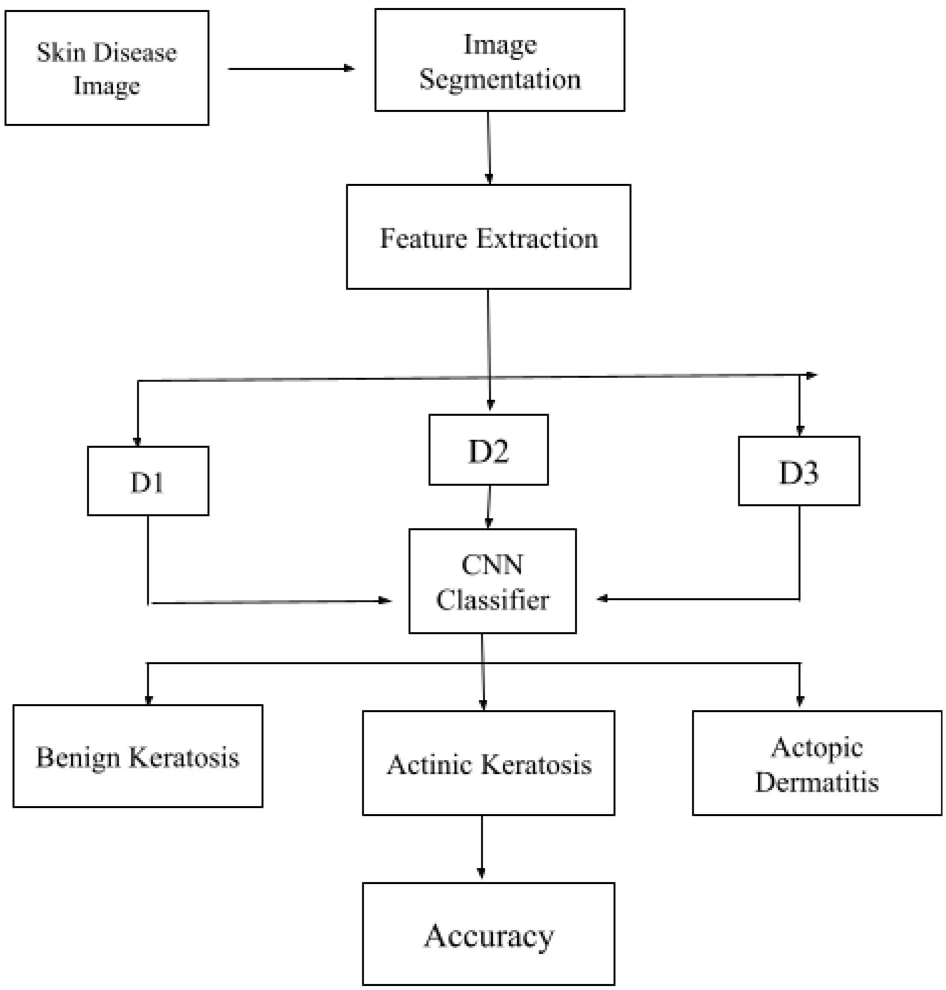
4) **Using Artificial Intelligence as a Diagnostic Decision Support Tool in Skin Disease: Protocol for an Observational Prospective Cohort Study**

Anna Escalé-Besa; Aïna Fuster-Casanovas ; Alexander Börve ; Oriol Yélamos; Xavier Fustà-Novell ; Mireia Esquius Rafat; Francesc X Marin-Gomez ; Josep Vidal-Alaball

This study aims to perform a prospective validation of an image analysis ML model as a diagnostic decision support tool for the diagnosis of dermatological conditions. The validation dataset includes 100 cases with GP, teledermatology, ML model (top 5 diagnoses), and face-to-face dermatologist assessments. The ML model's precision, sensitivity, specificity, and accuracy for 44 skin disease classes will be determined using a confusion matrix. Binary classification will be applied for each class, with area under the curve calculated using one-versus-all methodology. Macro- and micro-averaging will address infrequent classes, while precision, recall, and F-measure will be calculated and averaged. Additionally, accuracy of the ML model's top 3 diagnoses will be assessed. It has to be considered that even if the ML model does not provide a better diagnosis than the doctor’s, it is expected to help the practitioner consider other differential diagnoses.

Research Methodology for the Project

1. Methods



Where
D1 - Train dataset
D2 - Test dataset
D3 - Validation dataset

2. Data Collection

- **Data Source and Preparation**

  The data utilized for this project was sourced from a dataset comprising images of various skin diseases, collected from multiple repositories including Kaggle and Google

Scholar. To prepare the data for predictive modeling, we organized the images into two primary folders: 'val' and 'train' .

This organization facilitated efficient data management and upload to a designated drive, ensuring accessibility and readiness for use in the prediction process.

3. Model Development

- **AI/ML Techniques**: Utilize deep learning algorithms like Convolutional Neural Networks (CNNs) for image classification.
- **Model Architecture**: Design the neural network architecture optimized for skin disease recognition.
- **Software and Tools**: Use frameworks such as TensorFlow, PyTorch, or Keras for model development.

4. Training and Validation

- **Dataset Splitting**: Divide the dataset into training, validation, and test sets.
- **Training Process**: Train the model using the training dataset and validate performance using the validation set.
- **Hyperparameter Tuning**: Optimize model parameters to improve accuracy and generalization.

5. Evaluation Metrics

- **Performance Measures**: Assess model performance using metrics like accuracy, precision, recall, and percentage.
- **Cross-Validation**: Implement cross-validation techniques to ensure there is no or less error and prevent overfitting.

6. Deployment and User Interface (Future Approach)

- **System Integration**: Develop an interface for users (patients or clinicians) to upload **images and receive diagnostic results.**
- **Real-Time Inference**: Implement the trained model for real-time inference on new image submissions.

7. Results Analysis and Interpretation (Future Approach)

- **Interpretation:** Analyze model predictions and provide insights into the effectiveness of AI in dermatological diagnosis and some pretreatment options or self care related to the disease.
- **Discussion**: Discuss the implications of findings and potential applications in clinical practice and how this model is user friendly and can be used effectively.

Data analysis & Interpretation

Dataset Description

We have used a dataset containing 60 photos belonging to 3 kinds of diseases i.e, Actinic keratosis, Actopic Dermatitis and Benign keratosis.
Source - https://www.kaggle.com/datasets/riyaelizashaju/skin-disease-classification-image-dataset/code

Dataset Preprocessing

This function, named `data_preprocessing`, prepares image data for skin disease classification. It rescales pixel values, applies random transformations like zooming and flipping, and generates batches of preprocessed images and their labels for training the model.

Model Architecture

1. Input Layer: Accepts RGB images with dimensions `(img_size[0], img_size[1], 3)`.

2. Convolutional Layers:
 - Four convolutional layers with increasing filter sizes, followed by batch normalization, max pooling, and dropout.

3. Pooling Layers: Max pooling with pool size `(2, 2)` after each convolutional layer.

4. Flatten Layer: Converts the output of the convolutional layers into a 1D array.

5. Fully Connected Layers:
 - Two dense layers with 256 units and specified activation function.
 - Batch normalization and dropout layers after each dense layer.

6. Output Layer: Dense layer with 3 units and softmax activation for classification into 3 classes.

7. Model Summary: Provides an overview of the model architecture and parameters.

Training Procedure

```
model.compile(optimizer='adam', loss='categorical_crossentropy', metrics=['accuracy', Precision(), Recall()])
history_model = model.fit(train_generator, epochs = 30, validation_data = validation_generator)

Epoch 1/30
6/6 [==============================] - 120s 20s/step - loss: 1.7319 - accuracy: 0.4333 - precision_1: 0.4483 - recall_1: 0.4333 - val_loss: 24.9213 - val_accuracy: 0.3333 -
Epoch 2/30
6/6 [==============================] - 119s 21s/step - loss: 1.0850 - accuracy: 0.6333 - precision_1: 0.6800 - recall_1: 0.5667 - val_loss: 108.1522 - val_accuracy: 0.3333
```

The model trains with the Adam optimizer, categorical cross-entropy loss, and metrics including accuracy, precision, and recall. Then, it trains the model for 30 epochs using training data provided by `train_generator`, validating performance on `validation_generator` after each epoch.

Evaluation and Interpretation

```
class_names = list(train_generator.class_indices.keys())

def predict_disease(model, img_path):
  img = tf.keras.utils.load_img(img_path, target_size=(img_size[0], img_size[1], 3), color_mode = 'rgb')
  array = tf.keras.utils.img_to_array(img)
  array = array / 255.0

  img_array = np.expand_dims(array, axis=0)
  preds = model.predict(img_array)

  #formatted_predictions = []
  for prediction in preds:
      formatted_predictions = [f'{value:.2f}' for value in prediction]

  top_prob_index = np.argmax(formatted_predictions)
  top_prob = round(float(formatted_predictions[top_prob_index].replace(",", "."))*100, 2)

  print("Probability for each class:", sorted(zip(class_names, formatted_predictions), key=lambda x: x[1], reverse=True))

  plt.imshow(tf.keras.utils.load_img(img_path, target_size=(img_size[0], img_size[1],3), color_mode = 'rgb'))
  plt.axis('off')
  plt.title(f"Class: {list(class_names)[top_prob_index]}; Prob: {top_prob}%")
  plt.show()
```

We have evaluated by predicting the class of a skin disease in an input image and displaying the prediction along with the probabilities for each class. This functionality can be useful for testing the performance of the trained model on new, unseen images to see how well it generalizes to real-world data.

Results

Here, the model has presented the probability of each class to which the images can belong.

```
[ ] predict_disease(model, "/content/drive/MyDrive/Data/val/Actinic keratosis/ISIC_0025803.jpg")

1/1 [==============================] - 1s 1s/step
Probability for each class: [('Actinic keratosis', '1.00'), ('Actopic Dermatitis', '0.00'), ('Benign keratosis', '0.00')]
```


```
/content/drive/MyDrive/Data/train/Actinipredict_disease(model, "c keratosis/ISIC_0024763.jpg")

1/1 [==============================] - 1s 908ms/step
Probability for each class: [('Benign keratosis', '0.91'), ('Actinic keratosis', '0.06'), ('Actopic Dermatitis', '0.02')]
```


```
1/1 [==============================] - 1s 1s/step
Probability for each class: [('Benign keratosis', '0.98'), ('Actinic keratosis', '0.02'), ('Actopic Dermatitis', '0.00')]
```

CONCLUSION

In conclusion, DermSelect stands as a revolutionary leap forward in the realm of skincare discovery. By harnessing the power of artificial intelligence and machine learning, DermSelect has transformed the daunting task of selecting skincare products into a personalized, empowering experience. Its ability to decode the intricacies of individual skin profiles with unparalleled precision marks a significant milestone in the ongoing quest for radiant, healthy skin.

Moreover, DermSelect's commitment to inclusivity, accessibility, and diversity ensures that every individual, regardless of their background or skin type, can embark on their skincare journey with confidence and empowerment. By embracing the complexities of skincare and offering tailored recommendations, DermSelect redefines beauty as a journey of self-discovery and self-care.

IMPLICATIONS

The implications of DermSelect extend far beyond the realm of skincare, permeating various sectors and industries. From a consumer perspective, DermSelect empowers individuals to make informed choices about their skincare routines, fostering a sense of agency and self-care. Moreover, DermSelect has the potential to revolutionize the beauty industry by optimizing product development and marketing strategies based on consumer feedback and preferences. Additionally, the integration of AI and machine learning in DermSelect underscores the transformative impact of technology on healthcare and wellness, paving the way for future innovations in personalized medicine and beyond.

SCOPE FOR FUTURE RESEARCH

1. **Refinement of Algorithms:** Future research could focus on further refining the algorithms and machine learning models used in DermSelect to improve the accuracy and precision of skincare product recommendations. This could involve exploring advanced techniques in deep learning, natural language processing, and computer vision to better analyze user-provided images and text data.
2. **Integration of Additional Data Sources:** Investigating the integration of additional data sources, such as genetic information, lifestyle factors, or environmental influences, could enrich the personalized skincare recommendations provided by DermSelect. Understanding how these factors interact with skincare products and impact individual skin health could lead to more comprehensive and tailored recommendations.

3. **Longitudinal Studies:** Conducting longitudinal studies to evaluate the long-term effectiveness and user satisfaction with DermSelect could provide valuable insights into its impact on skincare outcomes. Tracking users' skincare journeys over an extended period could help assess the sustainability of results and identify any potential challenges or areas for improvement.

4. **User Experience Optimization:** Future research could focus on optimizing the user experience of DermSelect, including interface design, usability, and accessibility considerations. Gathering feedback from users through surveys, interviews, and usability testing could inform iterative improvements to enhance user engagement and satisfaction.

5. **Personalized Skincare Regimens:** Exploring the development of personalized skincare regimens tailored to individual users' needs and goals could be an exciting avenue for future research. By combining data-driven recommendations from DermSelect with expert knowledge in dermatology and skincare science, researchers could design customized routines optimized for specific skin concerns and conditions.

6. **Validation and Certification:** Investigating the validation and certification of DermSelect as a reliable and trustworthy skincare resource could be essential for gaining broader acceptance and adoption. Collaborating with dermatologists, skincare experts, and regulatory bodies to validate the efficacy and safety of DermSelect's recommendations could help build trust among users and industry stakeholders.

7. **Expansion to Other Beauty Categories:** Considering the expansion of DermSelect beyond skincare to encompass other beauty categories, such as haircare, makeup, and personal care products, could broaden its utility and impact. Developing specialized algorithms and datasets tailored to these categories .

8. Future iterations of DermSelect could introduce a redesigned User Interface (UI) with a focus on intuitive navigation, streamlined workflows, and enhanced visual aesthetics. By incorporating principles of user-centered design and human-computer interaction, the new interface aims to elevate the user experience, making skincare discovery both effortless and enjoyable.

Overall, the scope for future research in DermSelect is vast and multifaceted, offering opportunities to advance skincare technology.

References

https://www.researchprotocols.org/2022/8/e37531/

https://onlinelibrary.wiley.com/doi/full/10.1002/ski2.81

https://www.researchgate.net/profile/Isaac-Nti-3/publication/337171348_A_Web-Based_Skin_Disease_Diagnosis_Using_Convolutional_Neural_Networks/links/5dc9b019a6fdcc575040ba16/A-Web-Based-Skin-Disease-Diagnosis-Using-Convolutional-Neural-Networks.pdf