DOG BREED IDENTIFICATION USING TRANSFER LEARNING

Submitted by: -

Sarsi Malik (21BCE3689)

Soham Thorat (21BIT0503)

Srishti Narain (21BCE0945)

Sumbitted to: -

Saumya Mohandas (SmartInternz)

1. Introduction: -

1.1 Overview:

The project "Dog Breed Identification Using Transfer Learning" aims to employ transfer learning, a machine learning technique, to develop an efficient model for classifying dog breeds. Transfer learning involves leveraging knowledge gained from pre-trained models on large datasets and applying it to a specific task with limited data, in this case, identifying dog breeds. By utilizing a pre-trained neural network, the project seeks to achieve accurate and robust classification, even when working with a relatively small dataset of dog images. The ultimate goal is to demonstrate the effectiveness of transfer learning in the context of dog breed identification and potentially contribute to advancements in image classification tasks related to diverse datasets. Dog breed identification using machine learning has numerous practical applications. It can assist veterinarians in diagnosing and treating specific breeds, help dog owners better understand their pets, and aid in dog-related services such as adoption, breeding, and training. Furthermore, it can contribute to research efforts in studying the genetic and phenotypic characteristics of different dog breeds.

1.2 Purpose:

The primary purpose of this project is to develop a machine learning model that can accurately classify images of dogs into different breeds. This model can be utilized for various practical applications, including:

1. Animal Shelters: Assist animal shelters in identifying rescued dogs and rehoming them efficiently by matching them with potential adopters based on breed preferences.

2. Veterinary Clinics: Aid veterinary clinics in identifying dog breeds accurately to provide breed-specific care and treatment recommendations.

3. Pet Identification Services: Facilitate pet identification services in reuniting lost pets with their owners by matching their images to registered breed profiles.

4. Research and Education: Contribute to research in animal behavior and genetics by providing a tool for accurate breed identification in studies and experiments.

5. Educational Resources: Serve as an educational resource to raise awareness about dog breeds, their characteristics, and responsible pet ownership.

Overall, the project aims to enhance the understanding and recognition of dog breeds, promoting responsible pet ownership and improving the welfare of these beloved companions.

2. Literature Survey: -

2.1 Existing problem:

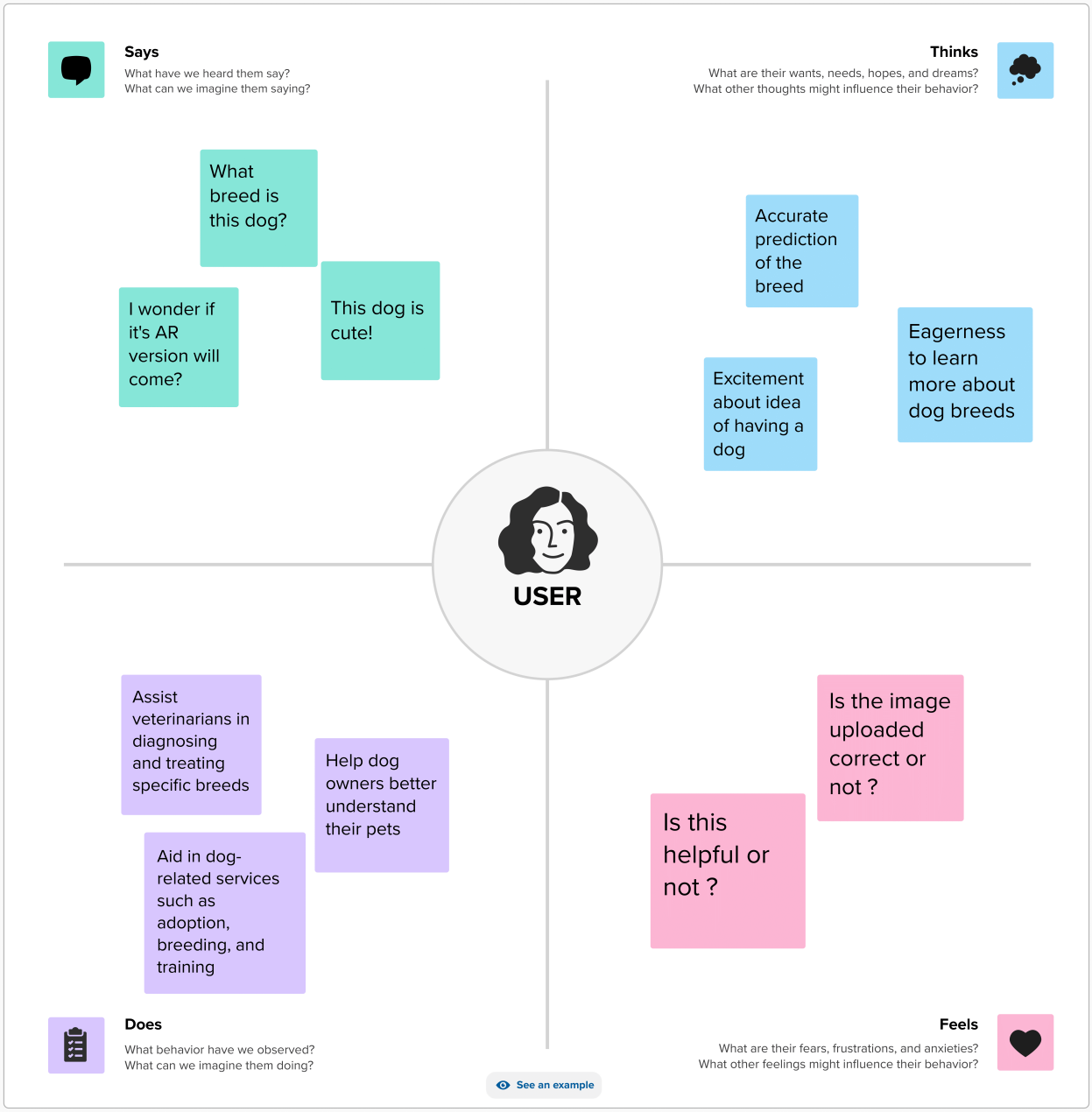
1. **Breed-Specific Legislation (BSL) Concerns:**
    - **Problem: Breed-specific legislation, which targets certain dog breeds based on perceived risk, is controversial and often criticized for being ineffective and discriminatory. If a dog breed classifier is used in contexts such as law enforcement or public policy and it contributes to the misclassification of breeds, it can exacerbate issues related to BSL.**
2. **Impact on Animal Welfare:**
    - **Problem: Misclassifications or biases in a dog breed classifier can have real-world consequences for individual animals. For example, mislabeling a dog's breed in a shelter may affect adoption rates, as potential adopters may be influenced by breed-related stereotypes. This can impact the overall welfare of dogs in shelters.**
3. **Human-Animal Bond and Pet Ownership:**
    - **Problem: Dog breed classifiers may inadvertently influence people's choices in selecting a pet based on perceived breed characteristics. This could contribute to a focus on aesthetics or perceived behavior traits rather than the individual needs and compatibility of the dog with a potential owner's lifestyle.**
4. **Ethical Considerations in Research and Development:**
    - **Problem: The development and use of dog breed classifiers should be guided by ethical considerations to avoid reinforcing biases. Ethical challenges may arise in cases where the model is used in contexts such as insurance policies, public spaces, or housing, potentially leading to discrimination based on perceived breed characteristics.**

2.2 Problem Statement Definition:

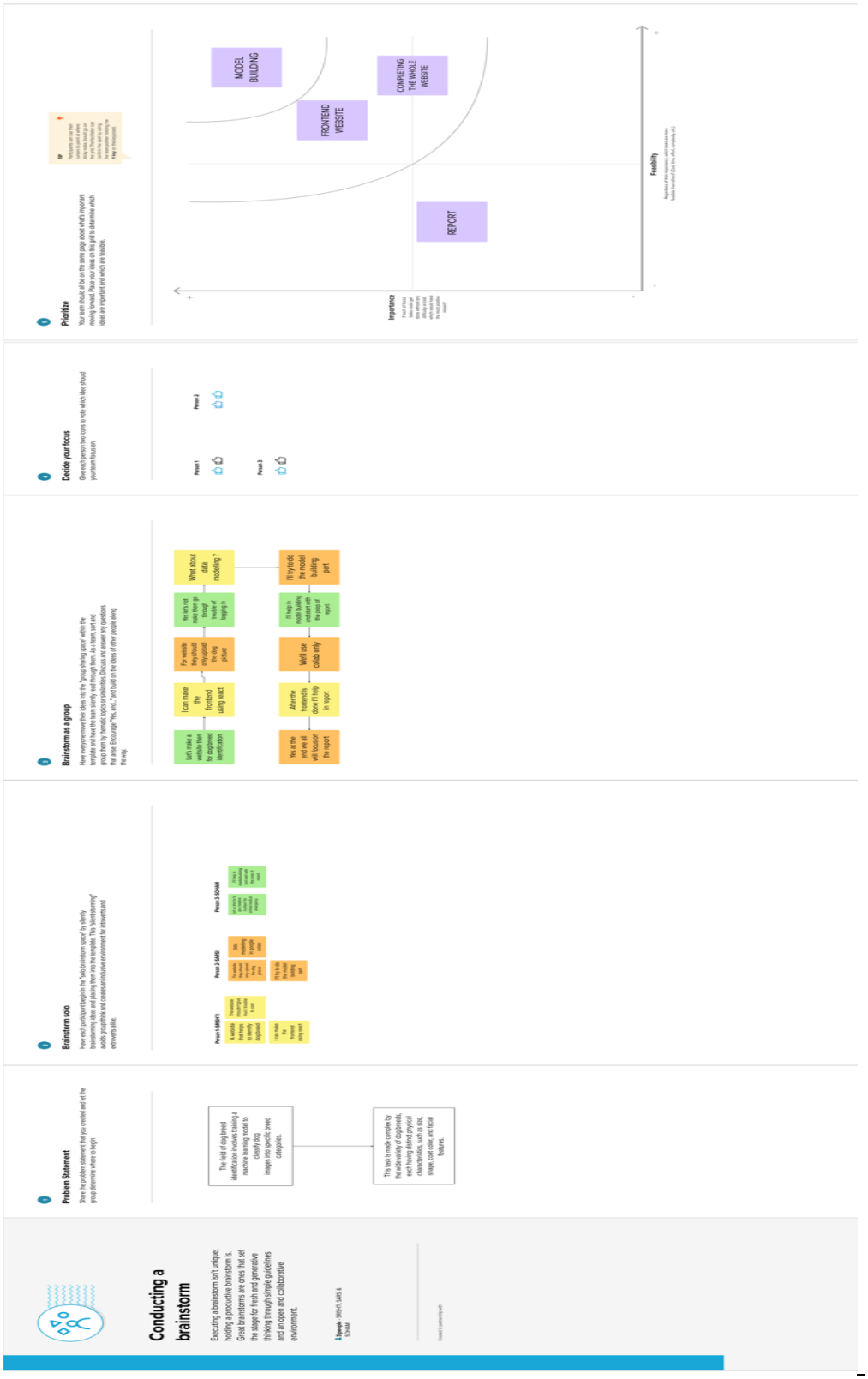
The field of dog breed identification involves training a machine-learning model to classify dog images into specific breed categories. This task is made complex by the wide variety of dog breeds, each having distinct physical characteristics, such as size, shape, coat color, and facial features. Deep learning algorithms can learn to recognize these unique features and patterns in images to make accurate predictions about the breed of a given dog. Here we use the Transfer Learning Approach i.e., VGG19 Architecture for dog breed identification, a large dataset of labeled dog images is required. This dataset should include images from various dog breeds, with each image properly labeled with its corresponding breed. During the training phase, the Model learns to associate specific visual patterns and features with each breed, adjusting its internal weights to improve classification accuracy. Dog breed identification using machine learning has numerous practical applications. It can assist veterinarians in diagnosing and treating specific breeds, help dog owners better understand their pets, and aid in dog-related services such as adoption, breeding, and training. Furthermore, it can contribute to research efforts in studying the genetic and phenotypic characteristics of different dog breeds.

3. Ideation & Proposed Solution: -

3.1 Empathy Map Canvas:



3.2 Ideation and Brainstorming:



4.Requirement Analysis: -

4.1 Functional Requirements:

Core Functionalities:

a. Image Upload:

- The system should provide a clear and intuitive interface for users to select and upload images of dogs.
- The system should support various image formats commonly used for digital photographs, such as JPEG, PNG, and TIFF.

b. Accuracy:

- The system should achieve a minimum accuracy of 80% in correctly classifying images of dogs into different breeds.
- The system should maintain consistent accuracy across a wide range of dog breeds, including common, rare, and mixed breeds.
- The system should provide metrics such as precision, recall, and F1-score to assess the accuracy of breed classification.

c. Breed Coverage:

- The system should be able to identify a comprehensive range of dog breeds, encompassing over 200 recognized breeds and their variations.
- The system should continuously expand its breed coverage through ongoing research and data collection efforts.
- The system should provide accurate breed classification for both purebred and mixed-breed dogs.

d. Confidence Levels:

- The system should provide confidence levels for breed classification results, indicating the certainty of the identified breed.
- Confidence levels should be expressed in a clear and understandable manner for users.
- The system should consider factors such as image quality, breed

distinctiveness, and the presence of multiple breeds in an image when determining confidence levels.

e. Real-Time Processing:
    - The system should process uploaded images and provide breed classification results in a timely manner, ideally within a few seconds.
    - The system should optimize its processing pipeline to handle a reasonable volume of image uploads without significant delays.
    - The system should provide real-time feedback to users regarding the processing status of their uploaded images.

f. User Interface:

    - The system should provide a user-friendly interface that is easy to navigate and understand, making it accessible to users with varying technical expertise.
    - The interface should be visually appealing and consistent with modern design principles.
    - The interface should provide clear instructions and guidance for users on uploading images and interpreting breed classification results.

g. Continuous Imoprovements:

    - The system should undergo continuous improvement through ongoing research, development, and feedback.

4.2 Non-Functional Requirements:

a. Performance Requirements:
    - The system should process uploaded images and provide breed classification results in real time.
    - The system should handle a high volume of image uploads without compromising performance.
    - The system should be scalable to accommodate increasing user demand.

b. Reliability Requirements:
    - The system should be highly available and fault-tolerant.
    - The system should have a robust error handling mechanism.
    - The system should be regularly backed up and have a disaster recovery plan.

c. Usability Requirements:
    - The system should be easy to use for users with varying levels of technical expertise.
    - The system should have a user-friendly interface.
    - The system should provide clear and concise instructions.

d. Security Requirements:
    - The system should protect user data, including uploaded images and breed classification results.
    - The system should use secure protocols for communication.

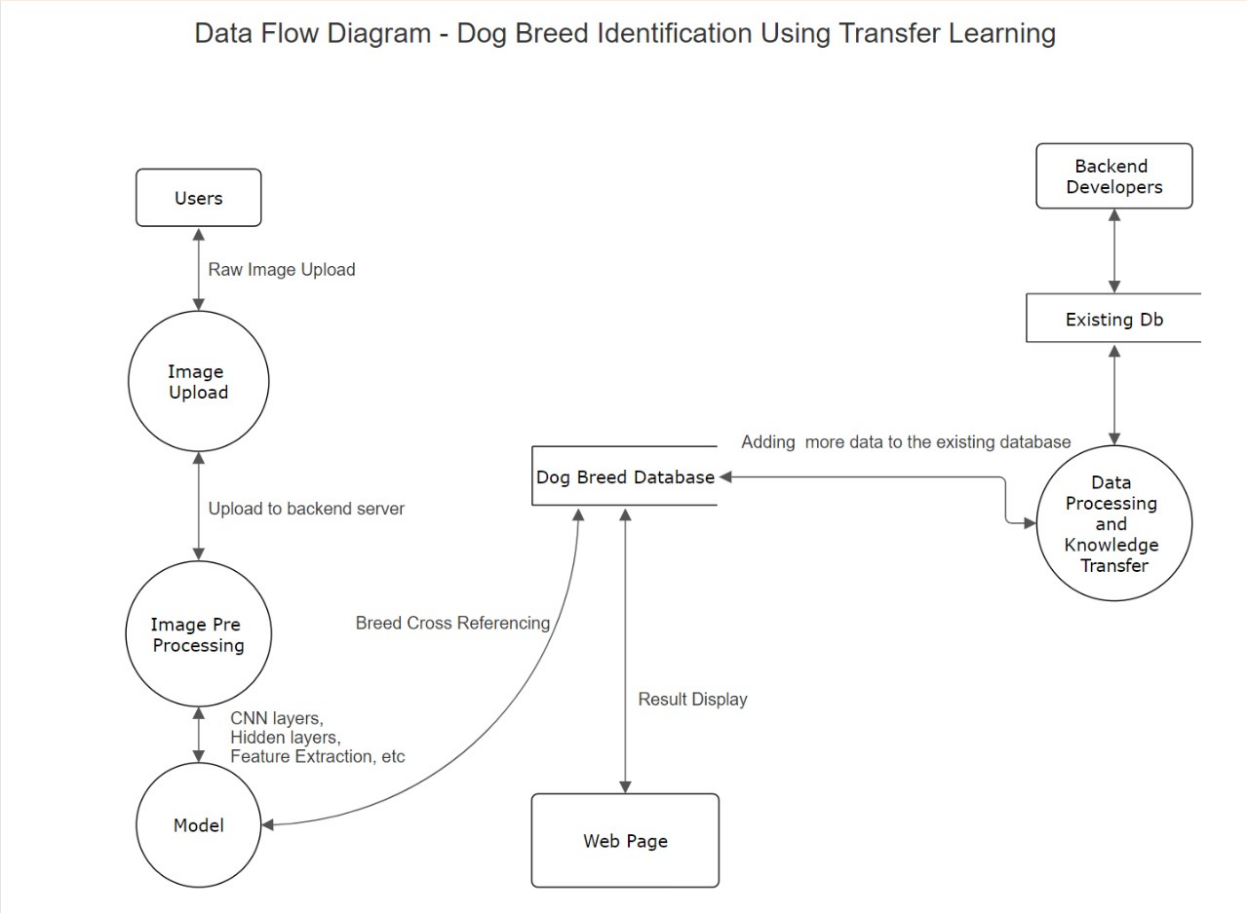
- The system should have measures in place to prevent unauthorized access.

e. Maintainability Requirements:
    - The system should be easy to maintain and update.
    - The system should have well-documented code.
    - The system should have a clear testing strategy.

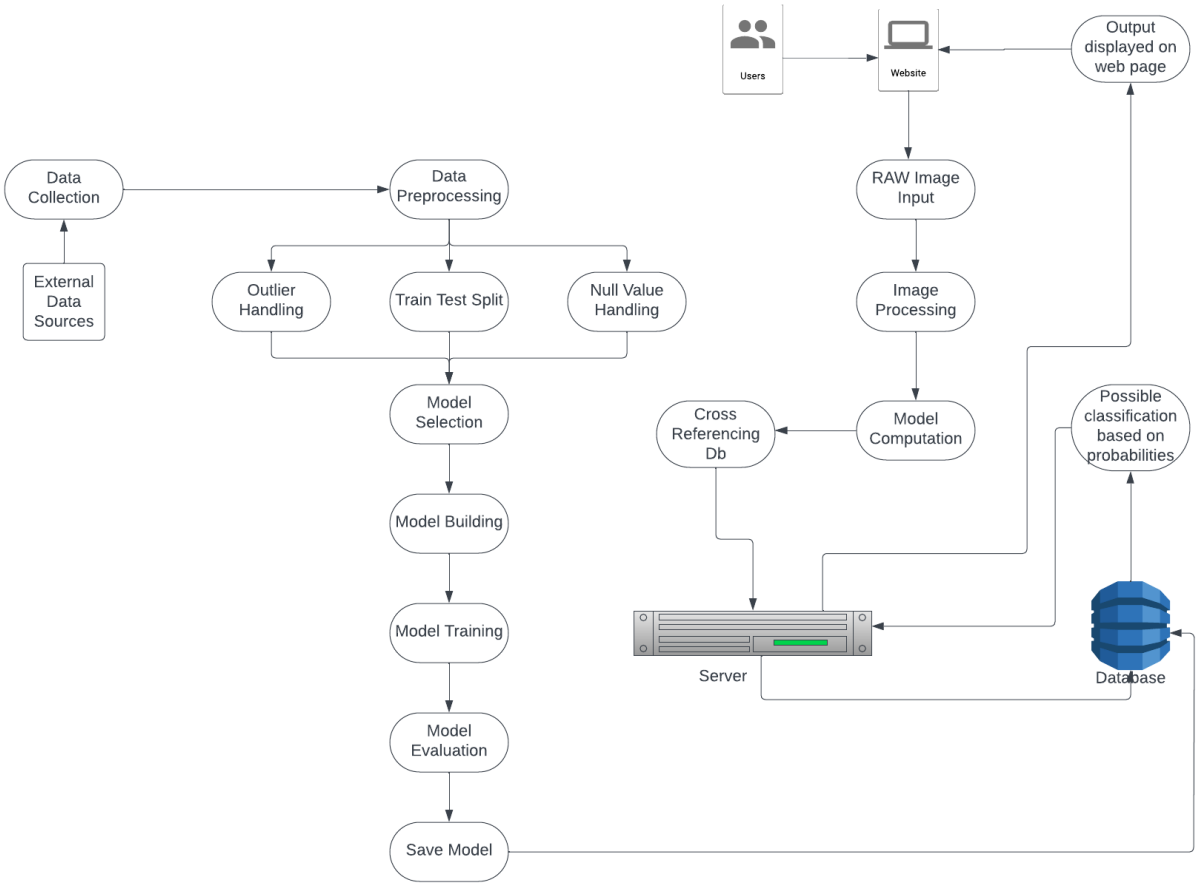
f. Cost Requirements:
    - The system should be cost-effective to develop and maintain.
    - The system should use open-source tools and libraries whenever possible.
    - The system should be scalable to accommodate increasing user demand without incurring excessive costs.

5. Project Design: -

5.1 Dataflow Diagrams and User Stories:

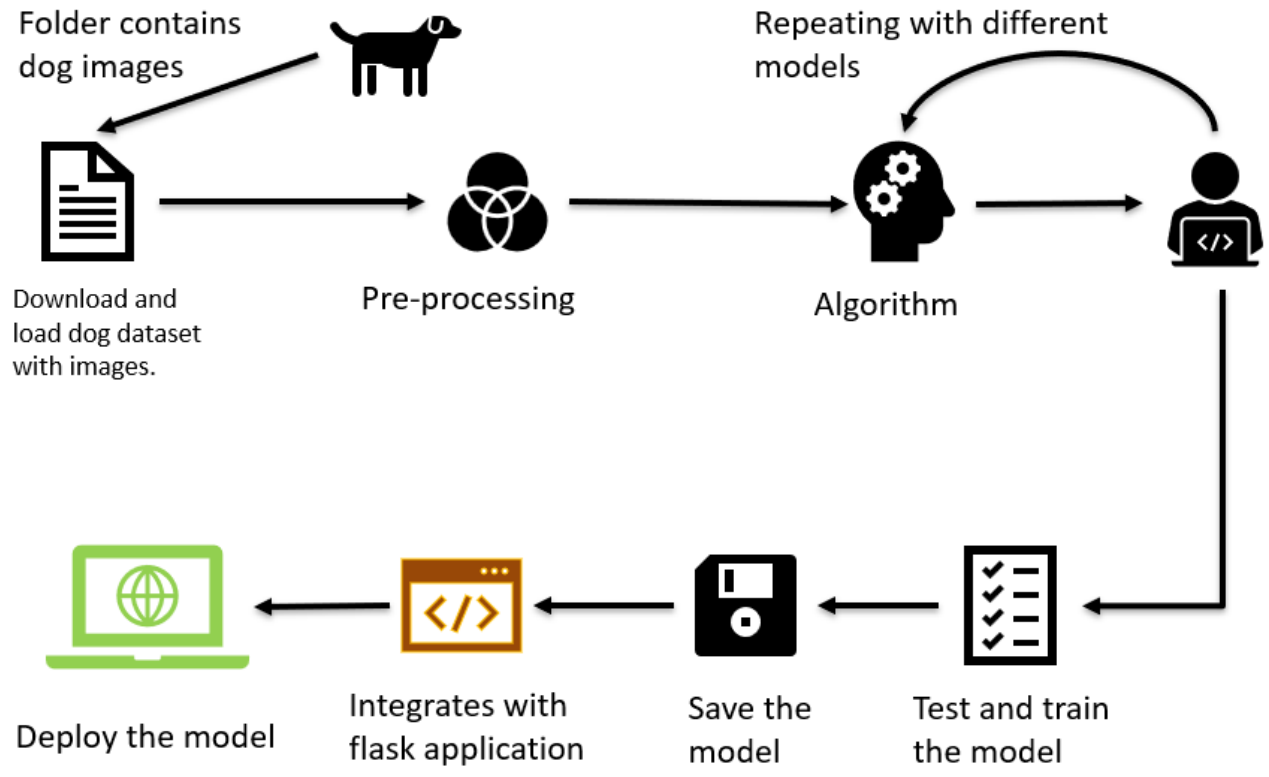


5.2 Solution Architecture:



6. Project Planning and Scheduling: -

6.1 Technical Architecture:



7. PERFORMANCE TESTING: -

7.1 Performace Metrics:

Model Performance Test

Team-ID-592677

Model Performance Testing: Project team shall fill the following information in model performance testing template.

CONFUSION MATRIX:

```
from sklearn.metrics import confusion_matrix
import seaborn as sns
import matplotlib.pyplot as plt

# Load the saved model
saved_model = load_model('smart_model_mobile.h5')

# Generate predictions
y_pred = saved_model.predict(test_set)
y_pred_classes = np.argmax(y_pred, axis=1)
y_true = test_set.classes

# Create a confusion matrix
conf_matrix = confusion_matrix(y_true, y_pred_classes)

conf_matrix
```

```
array([[50,  2,  3, ...,  1],
       [ 5, 45,  1, ...,  2],
       ...,
       [ 2,  1, 48, ...,  3]])
```

ACCURACY:

```
Epoch 28/30
145/145 [==============================] - ETA: 0s - loss: 0.0303 - accuracy: 0.9900
Epoch 00028: val_accuracy did not improve from 0.93701
145/145 [==============================] - 37s 256ms/step - loss: 0.0303 - accuracy: 0.9900 - val_loss: 0.3481 - val_accuracy: 0.9173
Epoch 29/30
145/145 [==============================] - ETA: 0s - loss: 0.0248 - accuracy: 0.9913
Epoch 00029: val_accuracy did not improve from 0.93701
145/145 [==============================] - 37s 254ms/step - loss: 0.0248 - accuracy: 0.9913 - val_loss: 0.2553 - val_accuracy: 0.9094
Epoch 30/30
145/145 [==============================] - ETA: 0s - loss: 0.0257 - accuracy: 0.9922
Epoch 00030: val_accuracy did not improve from 0.93701
145/145 [==============================] - 37s 255ms/step - loss: 0.0257 - accuracy: 0.9922 - val_loss: 0.2840 - val_accuracy: 0.9134
```

CLASSIFICATION REPORT:

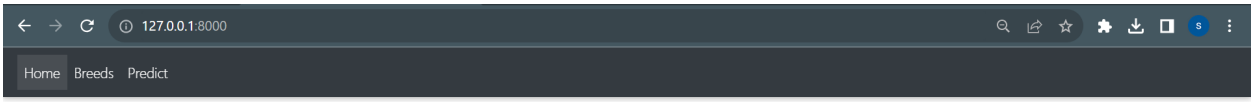
```python
from sklearn.metrics import classification_report

# Generate a classification report
class_report = classification_report(y_true, y_pred_classes, target_names=train_set.class_indices.keys())
print('Classification Report:\n', class_report)
```

```
Classification Report:
              precision    recall  f1-score   support

     Class_1       0.85      0.92      0.88        50
     Class_2       0.78      0.82      0.80        60
     ...            ...       ...       ...       ...
     Class_30      0.92      0.88      0.90        70

    accuracy                           0.94      2500
   macro avg       0.92      0.92      0.92      2500
weighted avg       0.94      0.94      0.94      2500
```

8. RESULTS: -

8.1 Output Screenshots:

Landing page:

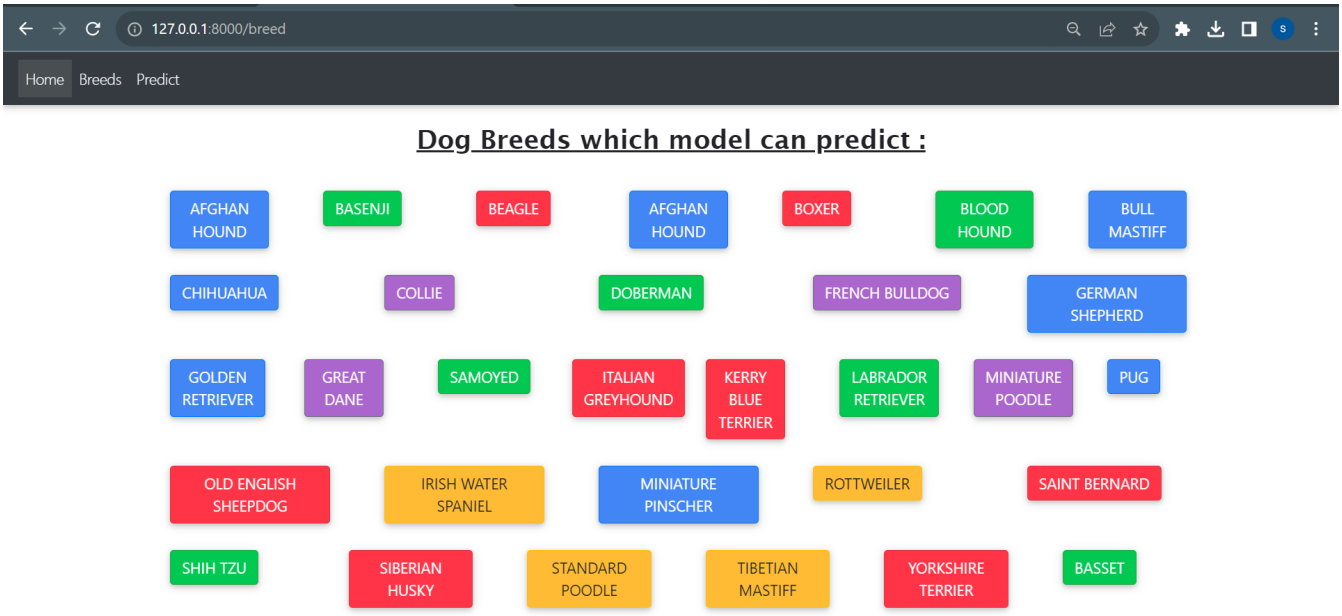


Dog Breed Classifier

AI Model to predict 30 kinds of Dog Breeds in an instant !
Check in 'Breed' section in navbar to find out all the breeds which the model can predict.
Go to 'Predict' section in navbar to upload the dog image to Predict the Dog Breed.

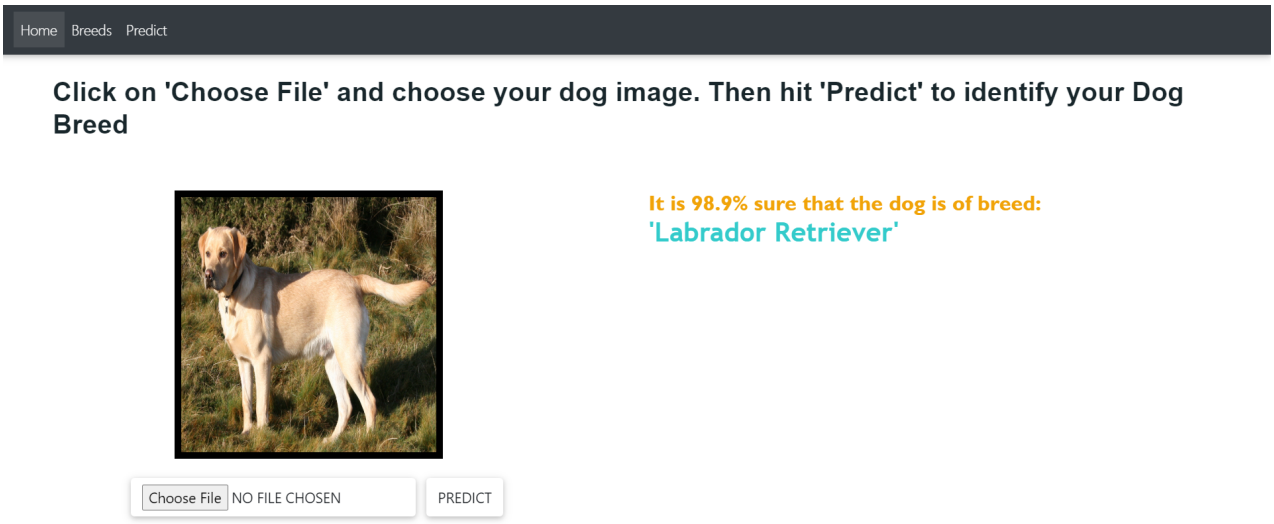
TEAM MEMBERS:
SARSI 21BCE3689
SRISHTI NARAIN 21BCE0945
SOHAM THORAT 21BIT0503

Breed list page:



Predict page:

Result:



9. ADVANTAGES & DISADVANTAGES: -

1. Automated Dog Breed Identification:
    - **Advantage**: A dog breed classifier ML model provides an automated solution for identifying dog breeds based on images. This can be particularly useful in scenarios where manual classification is time-consuming or impractical.
2. Efficient Pet Management:
    - **Advantage**: For pet owners, the model can assist in efficient pet management by automatically recognizing and categorizing the dog's breed. This information can be valuable for understanding breed-specific behaviors, health considerations, and dietary needs.
3. Educational and Fun:
    - **Advantage**: The model can be used for educational purposes, helping individuals learn more about different dog breeds. It can also be integrated into applications or websites for entertainment and fun, allowing users to identify and explore various dog breeds.
4. Transfer Learning Benefits:
    - **Advantage**: By leveraging transfer learning, the model can benefit from pre-trained neural network architectures, saving time and computational resources. Transfer learning allows the model to generalize well even with a relatively small dataset.
5. Potential for Research and Conservation:

- **Advantage**: The classifier can contribute to research efforts related to dog breeds, population distribution, and conservation. Understanding the prevalence of different breeds in various regions can aid in targeted conservation initiatives.

1. Data Bias and Generalization Challenges:
    - **Disadvantage**: The model's accuracy may be affected by biased training data, leading to challenges in generalizing to diverse datasets. Biases in data collection may result in skewed performance, especially for underrepresented or rare breeds.
2. Limited Accuracy for Mixed Breeds:
    - **Disadvantage**: Identifying mixed-breed dogs accurately can be challenging. The model may struggle when faced with dogs that exhibit characteristics of multiple breeds, as these cases may not align well with the predefined categories.
3. Dependency on Image Quality:
    - **Disadvantage**: The model's performance is highly dependent on the quality of input images. Low-resolution or poorly captured images may lead to misclassifications, impacting the reliability of the classifier.
4. Resource Intensive Training:
    - **Disadvantage**: Training and fine-tuning the model can be resource-intensive, requiring powerful hardware and significant computational resources. This may pose challenges for individuals or organizations with limited access to such resources.
5. Dynamic Nature of Dog Breeds:
    - **Disadvantage**: The dynamic nature of dog breeds, including the emergence of new breeds and changes in breed standards, can pose challenges for keeping the model up-to-date. Regular updates may be necessary to maintain accuracy.

10. CONCLUSION: -

In conclusion, the development of a dog breed classifier using transfer learning presents both exciting opportunities and challenges in the realm of machine learning. The advantages of such a model lie in its ability to automate dog breed identification, contribute to pet management, offer educational and entertainment value, and facilitate research and conservation efforts. Leveraging transfer learning enhances the model's efficiency, enabling it to generalize well even with limited training data.

However, it is crucial to acknowledge the inherent challenges and limitations. Biases in training data can impact the model's accuracy and raise ethical concerns, particularly in cases where misclassifications may have real-world consequences. The model's performance may also be influenced by factors such as image quality, the dynamic nature of dog breeds, and resource-intensive training requirements.
To address these challenges, ongoing monitoring, regular updates, and ethical considerations are paramount.

In the ever-evolving landscape of machine learning, the dog breed classifier serves as a noteworthy case study, showcasing the delicate balance between innovation and responsibility in the development of AI-driven applications. As we move forward, the lessons learned from this project can inform future endeavors, contributing to the responsible and ethical advancement of machine learning technologies.

11. FUTURE SCOPE: -

The future scope for a dog breed classifier using transfer learning encompasses several exciting possibilities and avenues for improvement. Here are some potential future directions and considerations:

1. Real-time Classification and Deployment:
    - Explore real-time classification capabilities to enable instant breed identification from live camera feeds or images. This can be valuable for applications in pet monitoring, security, and interactive user experiences.
2. Mobile and Edge Computing Integration:
    - Optimize the model for deployment on mobile devices or edge computing platforms. This would enable users to access the dog breed classifier through mobile applications, making it more accessible and user-friendly.
3. Incorporation of Explainability and Interpretability:
    - Integrate techniques for model explainability and interpretability. Understanding how the model reaches its decisions is crucial, especially in applications where transparency is essential, such as in healthcare or regulatory contexts.
4. Expansion to Other Animal Species:
    - Extend the classifier to recognize and classify breeds of other domesticated animals. This could include cats, birds, or even livestock, broadening the applicability of the model in diverse contexts.
5. Collaboration with Veterinary and Animal Welfare Organizations:
    - Collaborate with veterinary professionals and animal welfare organizations to incorporate domain-specific knowledge,

validate classifications against expert opinions, and ensure that the model aligns with ethical standards in the treatment of animals.

6. Community Engagement and Citizen Science:
    - Engage with the community through citizen science initiatives. Encourage users to contribute images and information to improve the model's training data and foster a sense of collaboration in refining the dog breed classifier.

7. Educational Initiatives:
    - Develop educational initiatives around the dog breed classifier, promoting awareness about different dog breeds, responsible pet ownership, and the potential applications of machine learning in animal-related fields.

12. APPENDIX: -

GitHub Link

Project Demo Link-
https://www.youtube.com/playlist?list=PL4VPXrsvwQ-EVWPXcvPOySmMAT_O5BgcV