WE ARE going to start with title name and then the existing models literature survey one by one

AKANKSHA will start –

Good morning everyone as we know that our project predicting lumpy skin disease has existing models and based on that our literature survey study is :

1ST RIYA :

Existing model ---- and will continue the given script

2nd ANKIT :

SAME AS RIYA WILL START as the second existing model is ---- the script

3rd SUYASH :

4th AKANKSHA

MOVING TOWARDS OUR PROPOSED MODEL OUR METHODOLOGY IS : the script riya

Then we are going to explain methodology – RIYA

ALGORITHMS – ANKIT ppt points

SYSTEM ARCHITECTURE – SUYASH ppt diagram in detail

UML – AKANKSHA ppt diagram with its purpose to use

END !

This is how we will present it tomorrow!

RIYA

The Literature Survey for the existing model "Lumpy Skin Disease Detector" is that

- The study explores the use of machine learning, specifically Convolutional Neural Networks (CNNs), to detect Lumpy Skin Disease (LSD) in cattle.
- Utilizes the TensorFlow framework to develop a robust model capable of early and accurate identification of LSD from images of cattle(Lumpy_Skin_Disease_Dete…).

2. Methodology

- A dataset of images, consisting of both healthy and LSD-infected cattle, is prepared. Preprocessing steps are applied to extract important visual features.
- A CNN model is constructed within the TensorFlow environment, using convolutional layers for feature extraction and pooling layers to reduce data dimensionality.
- The model is trained and validated on a diverse set of images to ensure adaptability and reliability under real-world conditions (Lumpy_Skin_Disease_Dete…).

3. Algorithms Used

- **Convolutional Neural Networks (CNNs)**: Utilizes convolutional layers to detect patterns in images, making it effective for distinguishing healthy from infected cattle.
- **TensorFlow**: A powerful deep learning framework that manages data flow and optimizes computations, allowing for efficient processing of large datasets.

4. Drawbacks

- **Dataset Limitation**: The current dataset is small, focusing mainly on northern Indian cattle, which may limit the model's generalizability to other regions.
- **Accuracy Variations**: Initial model accuracy was around 60% due to limited data. Further research and a larger dataset improved the accuracy to over 82%, showing the importance of diverse data(Lumpy_Skin_Disease_Dete…).

5. We are going to come with solutions like :

- **Expand Data Collection**: A larger and more varied dataset from different regions and cattle breeds would improve the model's accuracy and robustness.
- **Transfer Learning**: Using pre-trained CNN and as well as ANN models could accelerate the training process and enhance accuracy with less data.
- **Mobile Deployment**: Developing a mobile app could provide farmers with an accessible tool for real-time LSD detection, aiding in faster disease management.

ANKIT

The Literature Survey for the existing model "An Explainable Machine Learning Model for Lumpy Skin Disease Occurrence Detection" is that :

1. Methodology

- The study uses **machine learning classifiers** such as Decision Tree, SVM, KNN to detect LSD.
- The Random Forest algorithm showed the best performance with an accuracy of **95.7%**, and SHAP (SHapley Additive exPlanations) was used to explain the predictions

3. Algorithms Used

- **Random Forest**:
- **KNN**:
- **SVM**: A supervised learning model that finds the optimal hyperplane for separating data points into classes.

4. Drawbacks

- **Data Imbalance**: The dataset is initially unbalanced, which can lead to biased predictions. The study used **SMOTE** (Synthetic Minority Over-sampling Technique) to balance the classes, but it may not always fully address complex imbalances.
- **Black Box Nature**: Some machine learning models, especially Neural Networks, lack interpretability, making it difficult to understand why certain predictions are made (An_Explainable_Machine_…).

5. We are going to come with solutions like :

- **Improving Model Interpretability**: Use Explainable AI (XAI) techniques like LIME (Local Interpretable Model-agnostic Explanations) to make complex models more understandable for non-technical users.
- **Incorporate Real-Time Data**: Integrating real-time geospatial data could improve the accuracy of predictions and allow for dynamic forecasting.
- **Expand Feature Set**: Include more diverse features such as animal health records, genetic data, and specific insect vector information to enhance the robustness of predictions.

SUYASH

The Literature Survey for the existing model "Lumpy Skin Disease Detection Using Deep Learning" is that

- The research focuses on detecting Lumpy Skin Disease (LSD) in cattle using deep learning techniques.
- Aims to accurately identify LSD using the **DenseNet169** architecture, a convolutional neural network, achieving a classification accuracy of **93.10%** (Lumpy_Skin_Disease_Dete…).

2. Methodology

- The study utilizes a dataset of images containing both normal and lumpy skin cases.
- The **DenseNet169 model** is used for feature extraction and classification. Dense layers are added for binary classification of the images (Lumpy Skin vs. Normal Skin).

3. Algorithms Used

- **DenseNet169**: A deep CNN known for its ability to extract complex features from images, making it suitable for high-resolution visual tasks like disease detection.
- **ImageDataGenerator**: Used for image augmentation, introducing variability in the training data, which helps the model generalize better to unseen cases.

4. Drawbacks

- **Limited Generalizability**: DenseNet169 may require large datasets for optimal performance, and smaller datasets may limit the model's ability to generalize well to diverse cattle populations.
- **High Computational Requirements**: Training deep learning models like DenseNet169 demands significant computational resources, which may be a constraint for wider adoption.

5. Future Scope & Solutions

- **Expand the Dataset**: Incorporating a larger and more diverse dataset, including different breeds and environments, can enhance model generalization.
- **Develop a Mobile App**: The research plans to create a mobile app for real-time LSD detection and provide treatment information to farmers and veterinarians.

- **Incorporate Multi-modal Data**: Future iterations could combine image data with additional information, such as geographical and environmental factors, to improve prediction accuracy.

AKANKSHA

So after the research and studying literature survey of various research papers of existing model like

Various studies have explored different machine learning techniques for LSD prediction.

Afshari Safavi et al. (2022) used the Extra Trees Classifier and Artificial Neural Networks (ANN), achieving 89% accuracy, but they focused mostly on meteorological data, suggesting future work could use other type of data too.

Kumar et al. (2023) found Random Forest to be the best performer (92.1% accuracy), though the model could be prone to overfitting. They suggested cross-validation with varied datasets for better results.

Singh et al. (2023) tested several classifiers, such as Logistic Regression and Random Forest, with all models showing high accuracy, though the increased complexity was noted as a drawback.

Ujjwal et al. (2022) also favored Random Forest (94.7% accuracy) but emphasized the need to test other models like Gradient Boosting for further validation.

Alemayehu et al. (2013) focused on risk assessments in cattle trade, but suggested combining qualitative and machine learning methods for better predictions.

We came to know that there are common issues or drawbacks in existing system and we aim to eliminate those and build a model which will have various type of data and dataset images to implement prediction on, we will use both CNN AND ANN with random forest to achieve higher accuracy and precision that is called as Model Ensemble: Use Stacking or Blending to combine multiple models for improved accuracy. To keep it simple and less complex will not overfit models. **Overfitting** occurs when a machine learning model learns the training data too well, capturing even the noise and irrelevant patterns. This results in high accuracy on training data but poor performance on unseen data because the model fails to generalize. It is often caused by overly complex models with too many parameters relative to the dataset size.

EVERYONE :

Confusion Matrix Components:

1. **True Positives (TP)**: Correctly predicted positive cases.
2. **True Negatives (TN)**: Correctly predicted negative cases.
3. **False Positives (FP)**: Incorrectly predicted positive cases (actual negatives predicted as positive).
4. **False Negatives (FN)**: Incorrectly predicted negative cases (actual positives predicted as negative).

Accuracy:

$$\text{Accuracy} = \frac{\text{Correct predictions}}{\text{All predictions}}$$

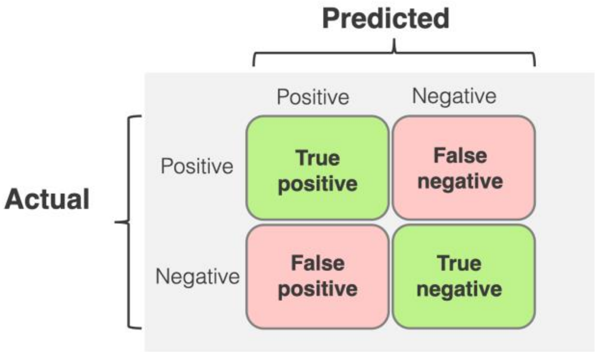
$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision: Important when the focus is on **minimizing false positives**, particularly in cases where predicting something incorrectly as positive has high consequences

Precision = True positives/ (True positives + False positives)

Confusion matrix :

A confusion matrix is a two-dimensional matrix used in classification experiments to evaluate the performance of a system by showing the number of correctly and wrongly classified data, helping to identify which classes of data are most often misplaced.


Predicted
Positive
Negative
Actual
Positive
True positive
False negative
Negative
False positive
True negative

DATASET :