**Indian Election Tweets Sentimental Analysis**

Submitted in the Partial Fulfilment of the Requirements for the Degree of

# BACHELORE OF TECHNOLOGY

## in

**COMPUTER SCIENCE & ENGINEERING**

**by**

Uday Sharma

**(210060101170)**

**Under the supervision of**

## Mr Vineet Kumar



Submitted to the

Department of Computer Science and Engineering

College of Engineering Roorkee (COER), Roorkee

Veer Madho Singh Bhandari Uttarakhand Technical University, Uttarakhand-248001

JANUARY-2023

# CANDIDATE’S DECLARATION

I hereby declare that the work presented in this project titled, “**Live lumpy virus Detection and Prediction**” submitted by me in the partial fulfilment of the requirement of the award of the degree of **Bachelor of Technology (B.Tech.)** submitted in the Department of **Computer Science & Engineering, Veer Madho Singh Bhandari Uttarakhand Technical University, Uttarakhand-248001**, is an authentic record of my thesis carried out under the guidance of **Dr Kamal Kant Verma, Head of Department of BCA , UETR**

**Date:** Uday Sharma

**B. Tech (CSE)**

**(210060101170)**

**College of Engineering Roorkee, Roorkee**

**Approved By Dr. Sumit Kumar**

**Head of the Department**

**(Computer Science & Engineering)**

**College of Engineering Roorkee, Roorkee**

# CERTIFICATE

It is to certify that the thesis entitled **“Live lumpy virus Detection and Prediction”** which is being submitted by **Uday Punia , Anshul Lakhera ,Yashas Gupta , Uday Sharma** to the Uttarakhand University Dehradun, in the fulfilment of the requirement for the award of the degree of **Bachelor of Technology (B. Tech.)** is a record of bonafide research work carried out by him under my guidance and supervision. The matter presented in this thesis has not been submitted either in part or full to any University or Institute for award of any degree.

**Dr Kamal Kant Verma**

**HOD of BCA ,UETR**

**College of Engineering Roorkee (COER), Roorkee**

Lumpy Virus Detection and Prediction

By- Uday Punia , Yashas Gupta , Anshul Lakhera and Uday Sharma



# Introduction

Lumpy skin disease virus (LSDV) causes an infectious disease in cattle. Due to its direct relationship with the survival of arthropod vectors, geospatial and climatic features play a vital role in the epidemiology of the disease. The objective of this study was to assess the ability of some machine learning algorithms to predict the occurrence of LSDV infection based on

meteorological and geological attributes. Initially, the ExtraTreesClassifier algorithm was used to select the important predictive features in forecasting the disease occurrence in unseen (test) data among meteorological, animal population density, dominant land cover, and elevation attributes. Some machine learning techniques revealed high accuracy in predicting the LSDV occurrence in test data (up to 97%). In terms of area under curve (AUC) and F1 performance metric scores, the artificial neural network (ANN) algorithm outperformed other machine learning methods in predicting the occurrence of LSDV infection in unseen data with the corresponding values of

0.97 and 0.94, respectively. Using this algorithm, the model consisted of all predictive features and the one which only included meteorological attributes as important features showed similar predictive performance. According to the findings of this research, ANN can be used to forecast the occurrence of LSDV infection with high precision using geospatial and meteorological

parameters. Applying the forecasting power of these methods could be a great help in conducting screening and awareness programs, as well as taking preventive measures like vaccination in

areas where the occurrence of LSDV infection is a high risk.



### Analysis tools

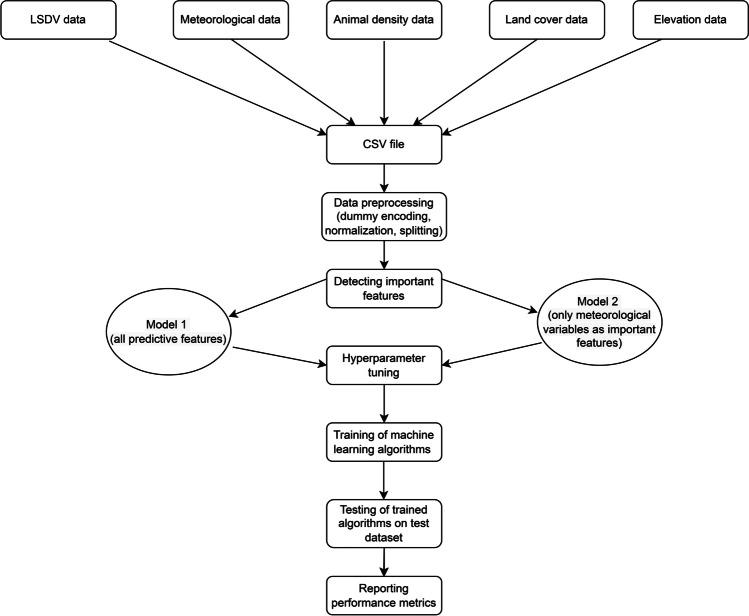
QGIS software (version 3.16 – Hannover) was applied to analyze and edit spatial data files. Machine learning techniques were implemented using the Python programming language (version 3.10) and the Anaconda navigator platform (as a package manager; version 1.10.0). Scikit-learn 0.24.1 (Pedregosa et al. [2011](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8759057/#CR40)) was used to implement logistic regression, SVM, decision tree, random forest, AdaBoost, and bagging algorithms. The XGBoost library was utilized to run the XGBoost technique (Chen and Guestrin [2016](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8759057/#CR12)). Keras API (Chollet [2018](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8759057/#CR14))

running as an abstraction layer on top of TensorFlow 2 framework (Abadi et al. [2016](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8759057/#CR1)) was used for building multilayer perceptrons (ANN). Streamlit is an open source app framework in Python language. It helps us create web apps for data science and machine learning in a short time. It is compatible with major Python libraries such as scikit-learn, Keras, PyTorch, SymPy(latex), NumPy, pandas, Matplotlib etc.

# Methodology

### Materials and methods

Figure [1](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8759057/figure/Fig1/) depicts the summary of steps taken in the materials and methods, and the details of each step are explained in the following sections.



[Fig. 1](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8759057/figure/Fig1/)

Summary of steps taken in the materials and methods section

**Libraries Used**

* Matplotlib.pyplot
* Urllib3

## Pandas/pandas\_datareader

* Notifier

## Bs4(Beautifysoup for code)

### Data sources

**Lumpy skin disease outbreak data** Geographic coordinates of Lumpy Skin Disease outbreaks were obtained from Global Animal Disease Information System of FAO (Food and Agriculture Organization) (<https://empres-i.review.fao.org/>/). Relevant information between January 2011

and March 2021 including the specific time of the outbreak and the longitude and latitude of the outbreak point were downloaded.

**Meteorological data** Monthly cloud cover (percentage), diurnal temperature range (degrees Celsius), frost day frequency (days per month), wet day frequency (days), potential

evapotranspiration (millimeters per day), precipitation (millimeters per months), daily mean temperature (degrees Celsius), monthly average maximum and minimum temperature (degrees Celsius), and vapor pressure (hectopascal) data for the period January 2011–December 2019

were obtained from the University of East Anglia’s Climatic Research Unit (CRU TS4.04) (Harris et al. [2020](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8759057/#CR27)).

**Animal density data** Cattle and buffalo population density data were obtained from Gridded Livestock of the World (GLW 3) database (Gilbert et al. [2018](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8759057/#CR23)).

**Land cover data** GLC-SHARE Beta-Release v1.0 (Latham et al. [2014](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8759057/#CR32)) was used to extract global land cover data (spatial information on various forms of physical covering of the Earth’s surface) including artificial surfaces, cropland, grassland, tree covered areas, shrubs covered

areas, herbaceous vegetation, aquatic or regularly flooded areas, mangroves, sparse vegetation, bare soil, snow and glaciers, and waterbodies coverage.

**Elevation data** Global geospatial elevation dataset (GRAY\_50M\_SR.VERSION 2.1.0) was downloaded from Natural Earth database (free vector and raster map data @ naturalearthdata.com).

# Project Status

 Downloaded the dataset of lumpy virus disease and its outbreak. - Completed  Learned about the particular libraries used in the project. - Completed

 Learned basic python and data science - Completed  Learned Machine learning

 part 1 - visualization and representation - Completed  Part 2 - prediction - In progress

# Conclusion & Future Scope

In this work, we used a statistical approach to identify major changes in the data underlying LSD outbreak reports. Additionally, we utilized time series models to forecast the number of LSD outbreak reports in Africa, Europe, and Asia during 2022–2024. Although LSD outbreak reports in Africa appear to be decreasing since 2020, it is expected that the number of reports will increase slightly. The number of LSD outbreak reports in Europe is projected to continue the previous 5-year steady trend. Additionally, the forecast predicts an increase in the number of outbreak reports in Asia. These findings indicate that LSD remains a substantial threat to the

cattle industry in various countries; thus, efforts should be made to monitor its spread within and between regions. Additionally, because LSD is regarded as a significant transboundary disease,

strict disease prevention and control in every country are critical. Furthermore, coordination among nations to control and eradicate the disease is essential.

# References

* Downloaded dataset of lumpy virus and skin disease from

[**https://data.mendley.com/datasets/7pyhbzb2n9/1**](https://data.mendley.com/datasets/7pyhbzb2n9/1)

* Found information about the disease and the machine learning algorithms used in it from

- [**https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8759057/**](http://www.ncbi.nlm.nih.gov/pmc/articles/PMC8759057/)

## ACKNOWLEDGEMENT

At this ecstatic time of presenting this project, first, the author bows to almighty God for blessing with enough patience and strength to go through this challenging phase of life. I would like to express a deep sense of gratitude and thanks to those people who have helped me in the accomplishment of this B. Tech. thesis.

First and foremost, I would like to thank my supervisor, **Name of the supervisor** for their expertise, guidance, enthusiasm, and patience. These were invaluable contributors whose insightful guidance helped to the successful completion of this project and spent many hours patiently answering questions and troubleshooting the problems.

Finally, I would like to thank all faculty, college management, administrative and technical staff of **College of Engineering Roorkee, Roorkee** and **Veer Madho Singh Bhandari Uttarakhand Technical University, Uttarakhand-248001** for their encouragement, assistance, and friendship throughout my candidature.

**Date:** Uday Punia ,Anshul Lakhera ,Yashas Gupta , Uday Sharma

**(210060101169, 210060101031, 210060101187, 210060101170)**

## PROJECT APPROVAL SHEET

This is to certify that the project titled

**Live lumpy virus Detection and Prediction**

By

Uday Punia , Anshul Lakhera ,Yashas Gupta , Uday Sharma

**Dr Kamal Kant Verma**

(HOD of BCA(UETR), College of Engineering Roorkee)