Data Collection and Preprocessing Phase

Date	10 July 2024
Team ID	739922
Project Title	Beyond The Veil Of Wellness: Machine Learning's Unique Journey In Animal Health Classification
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification Report:

Elevate your data strategy with the Data Collection plan and the Raw Data Sources report, ensuring meticulous data curation and integrity for informed decision-making in every analysis and decision-making endeavor.

Section	Description
Project Overview	"Beyond The Veil Of Wellness" leverages machine learning to enhance animal health classification and diagnosis. The project involves developing robust algorithms using comprehensive animal health data, including images and medical histories. These models will be integrated into a user-friendly platform for veterinarians, allowing continuous improvement based on new data and user feedback.
Data Collection Plan	Gather data from veterinary clinics, hospitals, research institutions, and public databases. Collect medical records, diagnostic images, behavioral data, and environmental information. Use electronic health records (EHR) integration for existing digital data and manual entry for paper records. Ensure comprehensive and diverse data for robust machine learning model training.

Raw Data Sources Identified	• **Veterinary Clinics and Hospitals**: Health records, diagnostic images, and treatment histories. • **Research Institutions**: Veterinary science databases. • **Public Databases**: Open-access animal health information and case studies. • **Animal Shelters and Zoos**: Health and behavioral records. • **Field Studies**: Data from wildlife research and conservation projects.

Data Collection Plan:

Source Name	Description	Location/URL

| Kaggle Dataset | The dataset comprises Animal name (symptoms1, symptoms2, symptoms3, symptoms4, Symptoms5). Dangerous. | https://www.kaggle.com/datasets/gracehephzibahm/animal-disease |