Paper/Journal 1

Title: Prediction and Detection of Liver Diseases using Machine Learning

Year and Author: Published in 2023 at the 3rd International Conference on Technology, Engineering, Management for Societal Impact using Marketing, Entrepreneurship, and Talent (TEMSMET). Authors: R.T. Umbare, Omkar Ashtekar, Aishwarya Nikhal, Bhagyashri Pagar, Omkar Zare.

Objective: The study aims to develop a machine learning-based model for early detection and classification of liver diseases. By leveraging ML techniques, the model can differentiate between healthy individuals and liver disease patients, further categorizing the type and severity of the disease while providing necessary precautions.

Dataset (Data Parameters and Sources):

- Dataset: Indian Liver Patient Dataset (ILPD) from UCI Machine Learning Repository.
- Features include:
    - Age, Gender
    - Total Bilirubin, Direct Bilirubin
    - Alkaline Phosphatase (ALP), Aspartate Aminotransferase (AST), Alanine Aminotransferase (ALT)
    - Total Proteins, Albumin, Albumin-Globulin Ratio
    - Diagnosis label (Liver disease or not)
- Data split:
    - **Training Set:** 80%
    - **Test Set:** 20%

Model Names:

1. **Support Vector Machine (SVM)**
2. **Logistic Regression (LR)**
3. **Naïve Bayes (NB)**
4. **K-Nearest Neighbors (KNN)**

Limitations:

- The dataset is relatively small and limited to a specific region, which may affect generalizability.
- The study does not explore deep learning models, which could improve accuracy.
- The model's effectiveness is dependent on feature selection and preprocessing.
- The research lacks real-time testing or clinical validation on real patient cases.

Research Gap:

- No comparison with deep learning techniques such as CNN or ANN.

- The study does not explore hybrid or ensemble models, which might yield better accuracy.
- Lacks integration with medical imaging data for a more comprehensive diagnosis.
- Requires further validation using external datasets from multiple geographic regions.

Result:

- The proposed models successfully classified liver disease with varying accuracy:
  - **SVM:** 73%
  - **Logistic Regression:** 73%
  - **Naïve Bayes:** 71.97%
  - **KNN:** 73%
- The models provide an efficient way to detect liver diseases and suggest necessary precautions.
- The research suggests that feature selection using genetic algorithms can improve prediction accuracy.