This thesis "**Machine Learning-Based Classification of Liver Steatosis Using Bioimpedance Spectral Features**" by **Noura Hanafy** (2025), investigates the use of bioelectrical impedance analysis (BIA) and machine learning (ML) to non-invasively detect and classify liver steatosis.

Problem Statement

Liver steatosis, or fatty liver disease, is an increasingly prevalent global health issue that can lead to severe complications like cirrhosis and carcinoma. While liver biopsy is the diagnostic "gold standard," it is invasive, time-consuming, and carries risks. Standard non-invasive methods like BIA often face limitations in accuracy due to physiological variability, low data variability, and the "curse of dimensionality" when handling high-dimensional spectral data.

Methodology

The research employed a quantitative

design to build a predictive model following a three-step workflow:

- **Input Data**: The study utilized Bioimpedance Spectrum (BIS) measurements consisting of 72 folders of experimental data. Data augmentation was applied to increase the small sample size (71 observations) and improve model robustness.
- **Feature Extraction & Dimensionality Reduction**: To manage the 30 initial features, various linear and nonlinear techniques were tested, including **Principal Component Analysis (PCA)**, **Kernel PCA**, **Singular Value Decomposition (SVD)**, **Autoencoders**, **t-SNE**, **Cole-Cole model**, and **Wavelet Transform**.
- **Classification**: Multiple nonlinear classifiers were evaluated to handle the complex, non-linear nature of the dataset, including **Nonlinear SVM**,

Random Forest, **k-Nearest Neighbors (kNN)**, **Naive Bayes**, **Quadratic Discriminant Analysis (QDA)**, **Ensemble Learners**, and **Deep Learning**.

Key Findings

- **Data Characteristics**: Initial ANOVA tests on raw data showed no significant difference between groups ($p > 0.05$), but **Z-score standardization** and **Min-Max normalization** successfully highlighted significant differences ($p < 0.05$), making the data more suitable for classification.
- **Model Performance**: The study compared the accuracy of different feature extraction-classifier combinations. Nonlinear dimensionality reduction techniques like **Kernel PCA** and the **Cole-Cole model** were expected to yield higher performance due to their ability to capture intrinsic data geometry.

- **Correlation**: Correlation matrices revealed that some techniques, such as Wavelet Transform, produced highly correlated features, which typically resulted in poorer classifier performance.

Conclusion

The thesis concludes that integrating ML with bioimpedance spectral features provides a promising, cost-effective, and non-invasive alternative for diagnosing liver steatosis. By utilizing standardization and advanced dimensionality reduction, ML models can effectively overcome the limitations of raw bioimpedance data to achieve reliable classification.