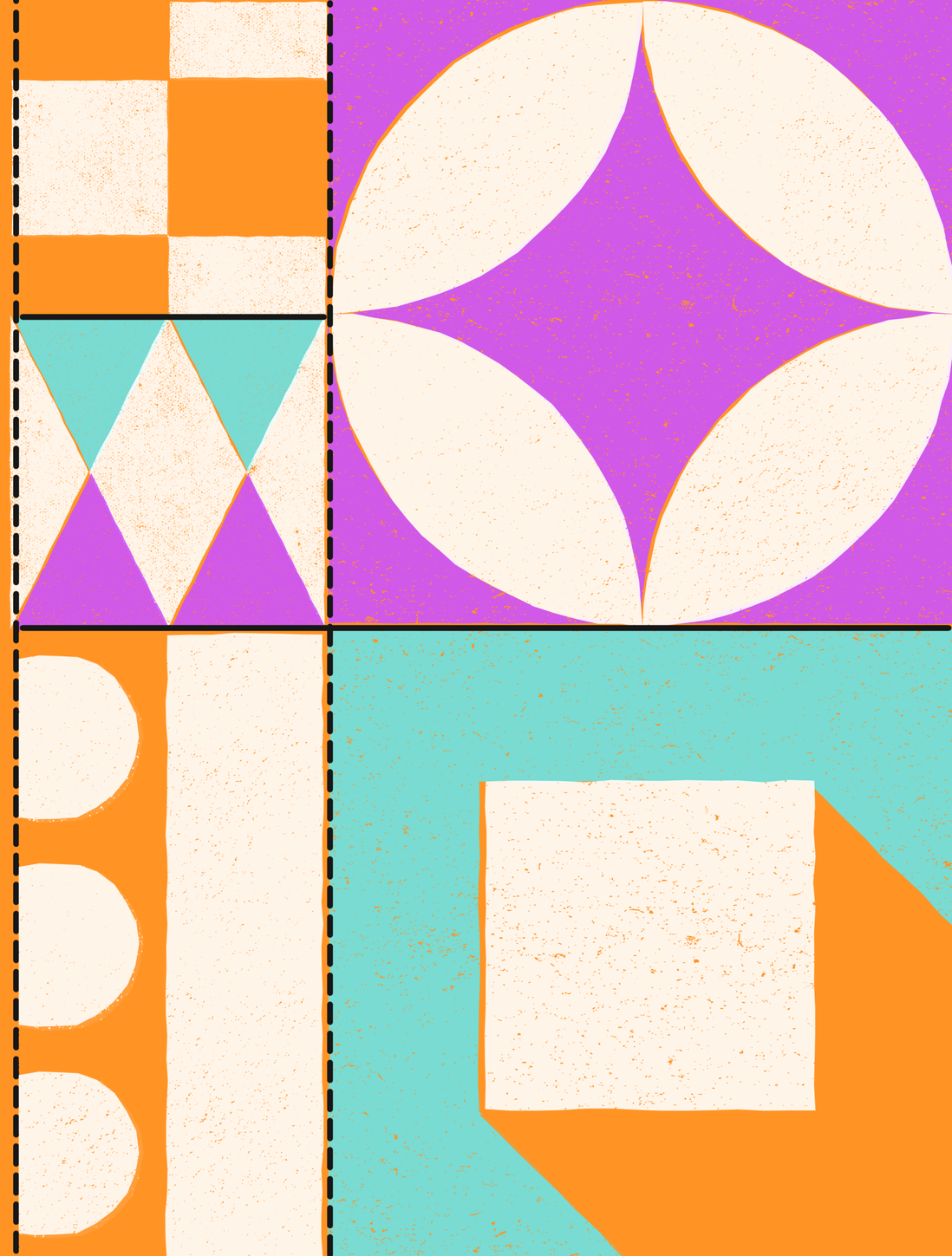
A survey paper on

Machine Learning Techniques to Identify Parkinson's Disease based on Different Types of Data

Submitted by
Aishwarya Barik,
Debarati Jash,
Souvik Ghosh,
Soumik Bannerjee, Poulami Mitra,
Chandra Das
and Shilpi Bose

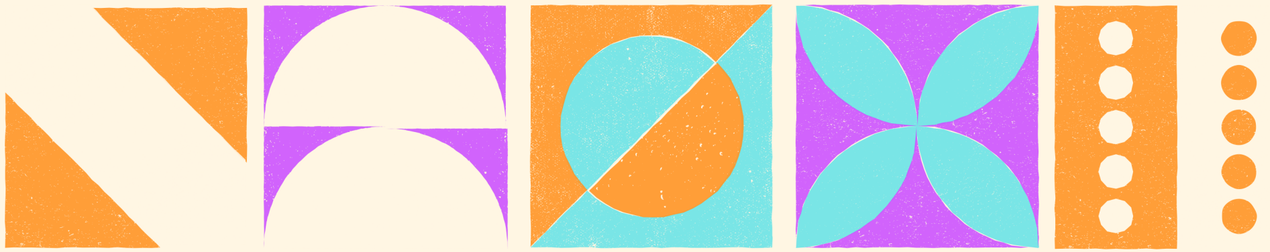
Netaji Subhash Engineering College



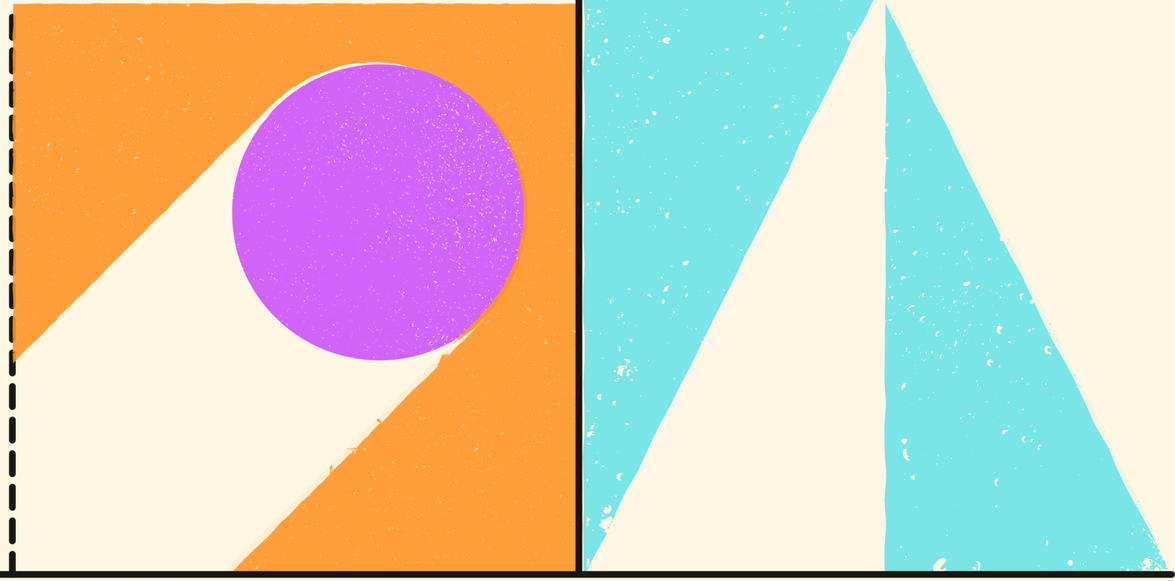
Introduction

This summary provides an overview of recent studies that have utilized machine learning methods to predict the progression of Parkinson's disease. The studies discussed in this summary cover a range of approaches, including the use of retinal imaging, MRI, and vocal features as input data. The results of these studies highlight the potential of machine learning models in accurately predicting the onset of neurodegenerative diseases and identifying individuals at risk. However, further research is necessary to validate these findings in larger cohorts and to gain a deeper understanding of the underlying biological mechanisms.

Authors :
Aishwarya Barik,
Debarati Jash,
Souvik Ghosh,
Soumik Bannerjee,
Poulami Mitra,
Chandra Das
and Shilpi Bose



Overview of Parkinson's Diseas



Parkinson's disease is a progressive neurological disorder that affects movement. It is caused by a loss of dopamine-producing neurons in the brain, particularly in a region called the basal ganglia, which plays a critical role in controlling movement.

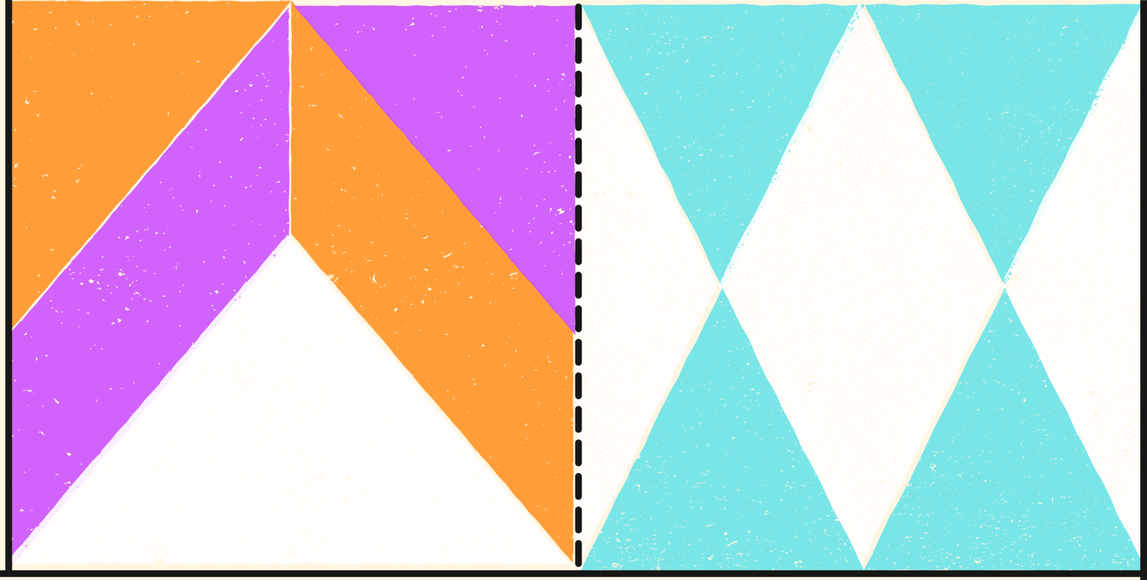
Symptoms: Tremors , Rigidity , Slow movements , Difficulty with balance and coordination

The symptoms of Parkinson's disease can vary in severity and can impact a person's daily activities and quality of life. Early detection of Parkinson's disease is important in order to provide appropriate treatment and support for those affected.

Impact on the nervous system:

- Reduced dopamine levels in the brain
- Changes in neural pathways and connections
- Inflammation and oxidative stress in the brain

Current Diagnostic Methods



Parkinson's disease is a complex disorder that can be difficult to diagnose accurately. Currently, there are several methods used for diagnosis, including clinical examination and imaging tests.

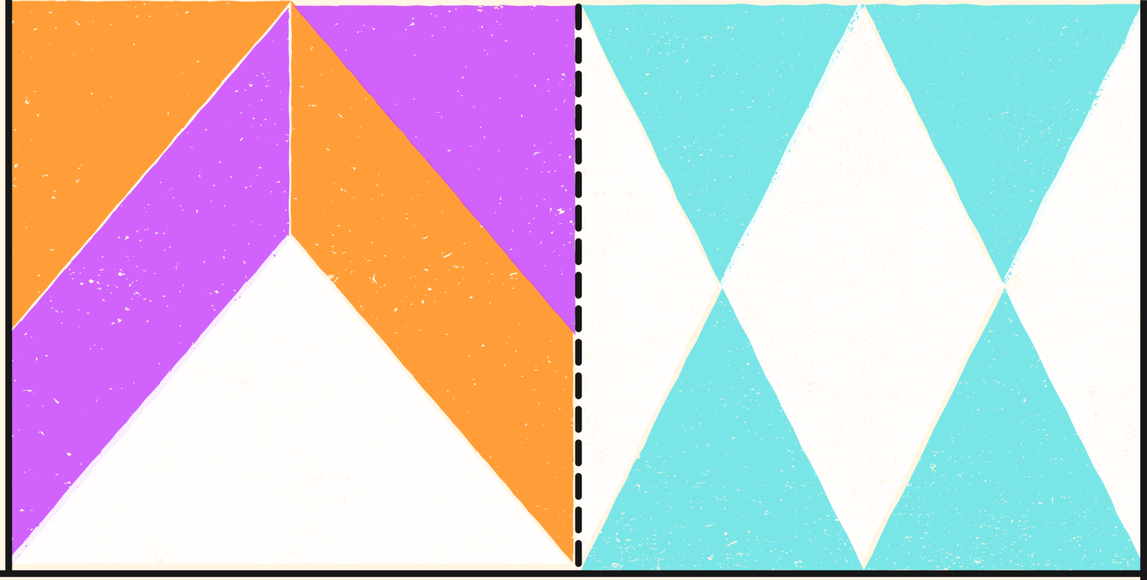
Clinical examination:

- A neurologist will conduct a physical examination and ask about medical history and symptoms.
- Pros: Non-invasive, does not require any special equipment.
- Cons: Subjective, may not detect early-stage Parkinson's disease.

Imaging tests:

- A range of imaging tests can be used to detect changes in the brain associated with Parkinson's disease, such as magnetic resonance imaging (MRI), computed tomography (CT), and dopamine transporter (DAT) imaging.
- Pros: Can detect changes in the brain associated with Parkinson's disease.
- Cons: Expensive, time-consuming, and not always covered by insurance.

Current Diagnostic Methods

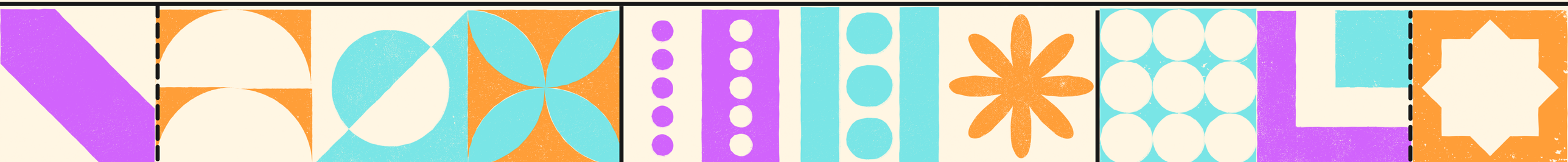


Diagnostic Method	Pros	Cons
Clinical examination	Non-invasive, does not require any special equipment	Subjective, may not detect early-stage Parkinson's disease
Imaging tests (e.g. MRI, CT, DAT imaging)	Can detect changes in the brain associated with Parkinson's disease	Expensive, time-consuming, and not always covered by insurance

Despite the advantages of these methods, there are limitations in their accuracy and accessibility, highlighting the need for alternative methods for early detection and diagnosis. Machine learning techniques have shown promising results in this regard, as they can analyze large datasets and identify patterns that may not be easily visible to the human eye.

Importance of Survey Paper on Parkinson's Detection with ML

- Parkinson's disease is a challenging neurological condition that affects a large number of people worldwide.
- There is a growing interest in using machine learning techniques for early detection and diagnosis of Parkinson's disease.
- A survey paper can help to provide a comprehensive overview of the existing research and highlight the most effective techniques and features used in Parkinson's detection with ML.
- Such a survey paper can help researchers and practitioners to stay up-to-date with the latest developments in the field, and guide future research directions.
- Additionally, a survey paper can provide insights into the limitations and challenges of current ML techniques, and identify areas where further research is needed.
- By synthesizing and organizing the current state of knowledge, a survey paper can help to accelerate progress towards more accurate and effective methods for detecting and diagnosing Parkinson's disease, ultimately leading to better patient outcomes.

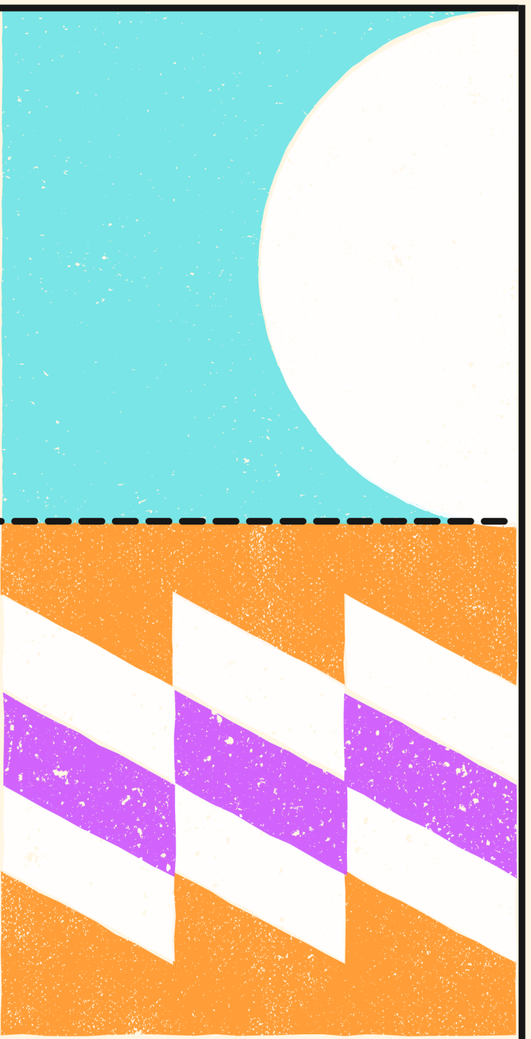


Types of Data Used in Machine Learning Techniques for Parkinson's Disease Detection

Parkinson's disease affects various aspects of human functioning, making it challenging to diagnose accurately. Machine learning techniques have shown promising results in detecting and diagnosing Parkinson's disease using various types of data. Some of the data types used in machine learning techniques for Parkinson's disease detection are:

- **Movement-related data:** This includes voice recordings that can detect changes in speech patterns, such as the hoarseness of voice, monotone speech, and speech rate changes.
- **Handwritten patterns:** Handwriting analysis can be used to detect changes in fine motor skills and writing speed, which are early signs of Parkinson's disease.
- **Multi-modal image data analysis:** Imaging techniques such as MRI, SPECT, and positron emission tomography are used to detect structural and functional changes in the brain.
- **Bio-molecular data analysis:** Proteins and other biomolecules in serum samples and cerebrospinal fluid can be analyzed to detect Parkinson's disease biomarkers.
- **Electromyography (EMG):** EMG measures the electrical activity of muscles and can be used to detect changes in muscle activity and motor control.
- **Optical coherence tomography (OCT):** OCT is a non-invasive imaging technique that can detect changes in the retina, which may be an early sign of Parkinson's disease.
- **Transcranial sonography (TCS):** TCS is an imaging technique that uses sound waves to create images of the brain and can detect changes in the basal ganglia, a brain region affected by Parkinson's disease.
- **Eye movements:** Eye movements can be analyzed to detect changes in saccades, smooth pursuit, and other eye movement patterns, which are early signs of Parkinson's disease.
- **Electroencephalography (EEG):** EEG measures the electrical activity of the brain and can detect changes in brain waves that may indicate Parkinson's disease.

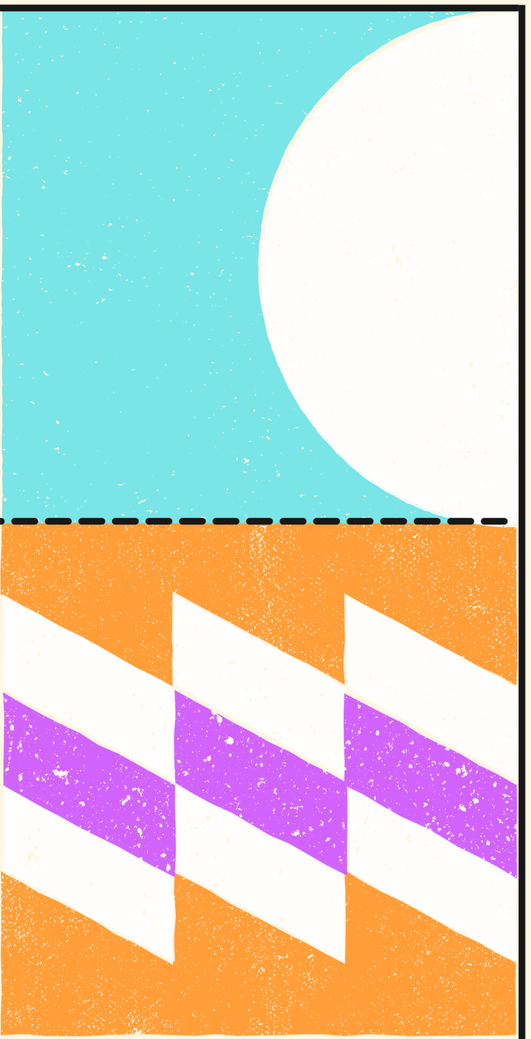
Overview of ML Techniques



Classification algorithms are used to predict the class of a data point based on its features

- **Artificial Neural Networks (ANN):** a type of deep learning algorithm that can learn complex relationships between features and classes
- **Support Vector Machines (SVM):** a binary classification algorithm that finds a hyperplane to separate the classes and maximizes the margin between them
- **Regression & Classification:** algorithms that can predict both continuous and categorical variables
- **NN classifier:** a type of neural network that can perform both regression and classification
- **Probabilistic Neural Network (PNN):** a type of neural network that uses Bayes' theorem to calculate the probability of a data point belonging to each class
- **Decision Tree:** a tree-based algorithm that makes a series of decisions based on the features of a data point to predict its class
- **Correlation based Feature Selection:** an algorithm that selects the most relevant features based on their correlation with the class variable
- **Multi-Layer Perceptron (MLP):** a type of neural network that can perform both regression and classification with multiple hidden layers
- **K-nearest neighbor (KNN):** an algorithm that predicts the class of a data point based on the classes of its nearest neighbors
- **Random Forest:** an ensemble of decision trees that can perform both regression and classification
- **Classification and Regression Trees (CART):** a type of decision tree that can perform both regression and classification
- **Adaboost.M1:** an ensemble algorithm that combines weak classifiers to form a strong classifier
- **Bagging:** an ensemble algorithm that trains multiple models on different subsets of the data and combines their predictions

Overview of ML Techniques



Clustering algorithms are used to group data points into clusters based on their similarity

- **Fuzzy C-means:** a soft clustering algorithm that allows data points to belong to multiple clusters
- **Vector Quantization:** a hard clustering algorithm that assigns each data point to a single cluster based on the closest centroid

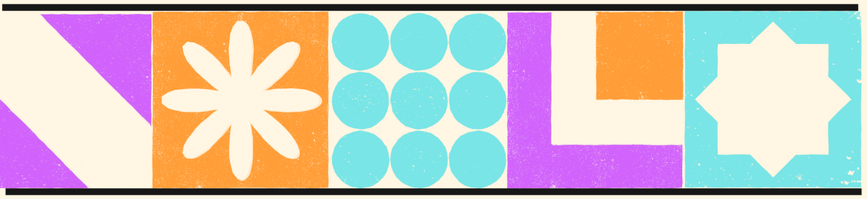
Feature selection algorithms are used to select the most relevant features for a predictive model

- **Correlation:** an algorithm that selects the most correlated features with the class variable
- **Entropy-based discretization method:** an algorithm that discretizes continuous features based on their entropy

Hybrid algorithms combine multiple techniques to improve predictive performance

- **Fuzzy k-nearest neighbor (FKNN):** a hybrid of fuzzy logic and KNN algorithm that assigns membership values to data points
- **Nested-SVM:** a hybrid of SVM and decision trees that uses decision trees to select the most relevant features and SVM to classify the data
- **Nonlinear dynamic:** a hybrid of neural networks and differential equations that can model complex nonlinear systems

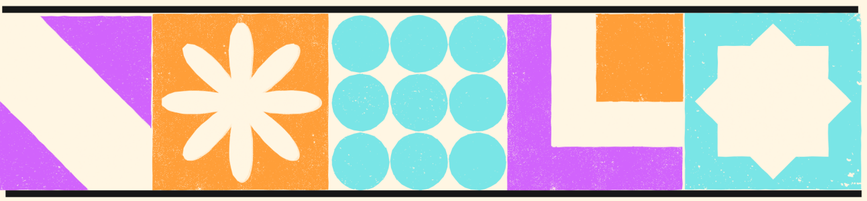
Paper Analysis 1



Enhanced vocal features in Parkinson's disease via machine learning
Author: Tsanas et al. (2021)

- Investigated vocal features of Parkinson's disease (PD) using machine learning
- Used a dataset of 103 participants, including 51 PD patients and 52 healthy controls
- Machine learning algorithms accurately distinguished between PD patients and healthy controls based on their vocal features
- Detected subtle changes in vocal features not visible to naked eye, useful for early diagnosis of PD
- Collected speech samples and extracted vocal features such as fundamental frequency, jitter, shimmer, and harmonics-to-noise ratio (HNR)
- Used k-nearest neighbors (KNN), support vector machine (SVM), decision trees, and artificial neural networks (ANN) for classification
- Proposed a novel feature extraction method based on empirical mode decomposition (EMD) for extracting features from the resulting intrinsic mode functions (IMFs)
- EMD features combined with traditional vocal features improved classification accuracy compared to using either method alone
- Proposed method has the potential to improve the accuracy of PD diagnosis and monitoring its progression.

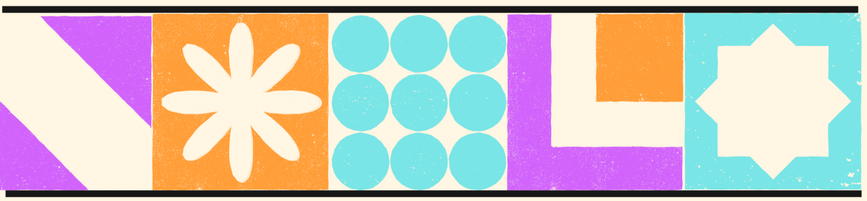
Paper Analysis 2



A gene expression-based prediction model for Parkinson's disease using machine learning
Author: Yang et al. (2021)

- The study aimed to develop a machine learning-based model for predicting PD using gene expression data.
- The dataset of 353 participants, including 193 PD patients and 160 healthy controls, was used in the study.
- Boruta was used to select the most relevant features (genes) that were differentially expressed between PD patients and healthy controls.
- A prediction model was developed using a random forest algorithm and clinical variables such as age and sex.
- The model was evaluated using metrics such as sensitivity, specificity, accuracy, and AUC-ROC and tested on an independent validation cohort.
- The machine learning models could accurately distinguish between PD patients and healthy controls with an accuracy of 91.20%.
- The study identified several genes that were significantly associated with PD, which could be useful for understanding the underlying mechanisms of the disease.

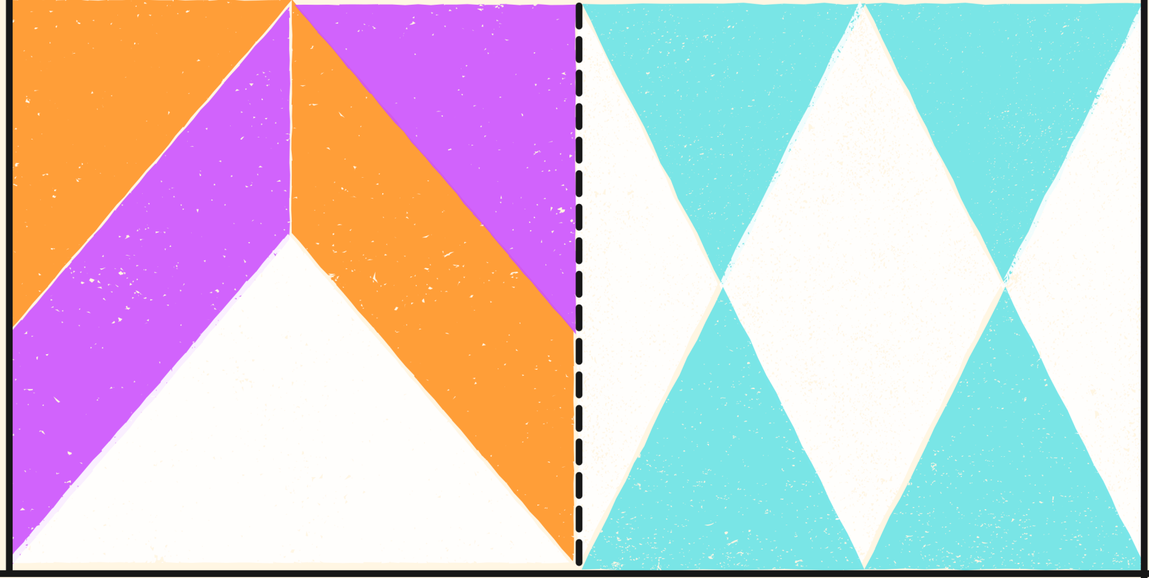
Paper Analysis 3



Diffusion magnetic resonance imaging-based machine learning for early diagnosis of Parkinson's disease
Author: Zhao et al. (year 2018)

- The study involved 98 participants, including 49 Parkinson's disease patients and 49 healthy controls.
- The diffusion MRI data were acquired using a 3.0 Tesla MRI scanner with 30 gradient directions and a b-value of 1000 s/mm2.
- FSL was used to preprocess the data, including skull stripping, eddy current correction, and tensor calculation to obtain DTI metrics.
- DTI metrics (FA, MD, AD, and RD) were used as features for a support vector machine (SVM) model to classify PD and HC participants.
- The SVM model was trained using leave-one-out cross-validation and achieved an accuracy of 95.3% in detecting Parkinson's disease.
- Statistical analysis was performed using the two-sample t-test to compare DTI metrics between PD and HC participants.
- The study demonstrates the potential of using advanced imaging and computational techniques for early diagnosis and treatment of Parkinson's disease.

Challenges and Opportunities



Challenges

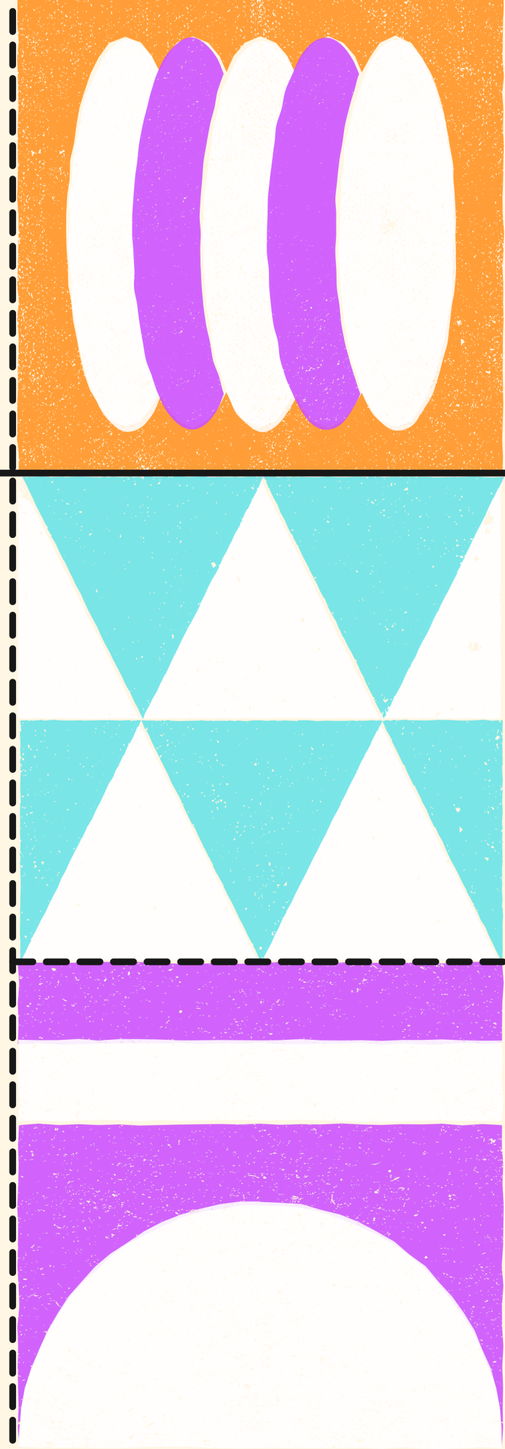
- Developing machine learning models that can accurately predict disease progression for individual patients based on their unique characteristics and genetic backgrounds.
- Integrating data from multiple sources, such as electronic health records and wearable devices, to provide a more comprehensive understanding of neurodegenerative diseases.

Opportunities

- Further research in this field will lead to earlier detection and more effective treatments for patients with these debilitating diseases.
- Advancements in technology and data collection will provide more data to feed machine learning algorithms, leading to more accurate predictions.

Overall, the potential of machine learning in predicting neurodegenerative diseases is enormous, but there are still many challenges and opportunities for future research. It is hoped that continued research in this field will lead to earlier detection and more effective treatments for patients with these debilitating diseases.

Conclusion



Parkinson's disease is a health concern with no known cure, but early detection and intervention can improve patient outcomes. Machine learning algorithms have shown great promise in predicting the onset and progression of these diseases using data from medical imaging, genomic, and clinical sources. The studies mentioned in this paper demonstrate the potential of machine learning in predicting neurodegenerative diseases and provide insights into the underlying mechanisms of these diseases.

Key Points

- Machine learning algorithms can accurately predict the onset and progression of Parkinson's disease using data from various sources.
- The potential of machine learning in predicting neurodegenerative diseases is enormous.
- Longitudinal studies that track disease progression over time are needed to provide a more accurate picture of disease progression and enable earlier detection.