

GROUP 7

ADVANCED PARKINSON'S DISEASE DETECTION WITH COMPREHENSIVE VOICE FEATURE ANALYSIS

2023DSS1013 - GAURAV SRIVASTAVA

2023DSS1014 - GEORGY S PULAYATH

2023DSS1023 - PRABHUJEET SAHU

2023DSS1032 - SHRENIK NANDRE

2023DSS1035 - SUMIT UPADHYAY



INTRODUCTION

What is Parkinson's Disease

Parkinson's Disease (PD) represents a significant difficulty within neurological disorders. A degenerative movement illness impacting the neurological system, Parkinson's Disease is characterized by debilitating symptoms including tremors, stiffness, bradykinesia, and postural instability.

Existing Issue

Due to the unavailability of exact diagnostic tests and the invasiveness of several current procedures, diagnosing PD creates complex issues. Blood testing, laboratory assessments, and brain imaging are employed to exclude alternative potential ailments, although these techniques may involve strenuous procedures that could exacerbate the discomfort of Parkinson's disease patients



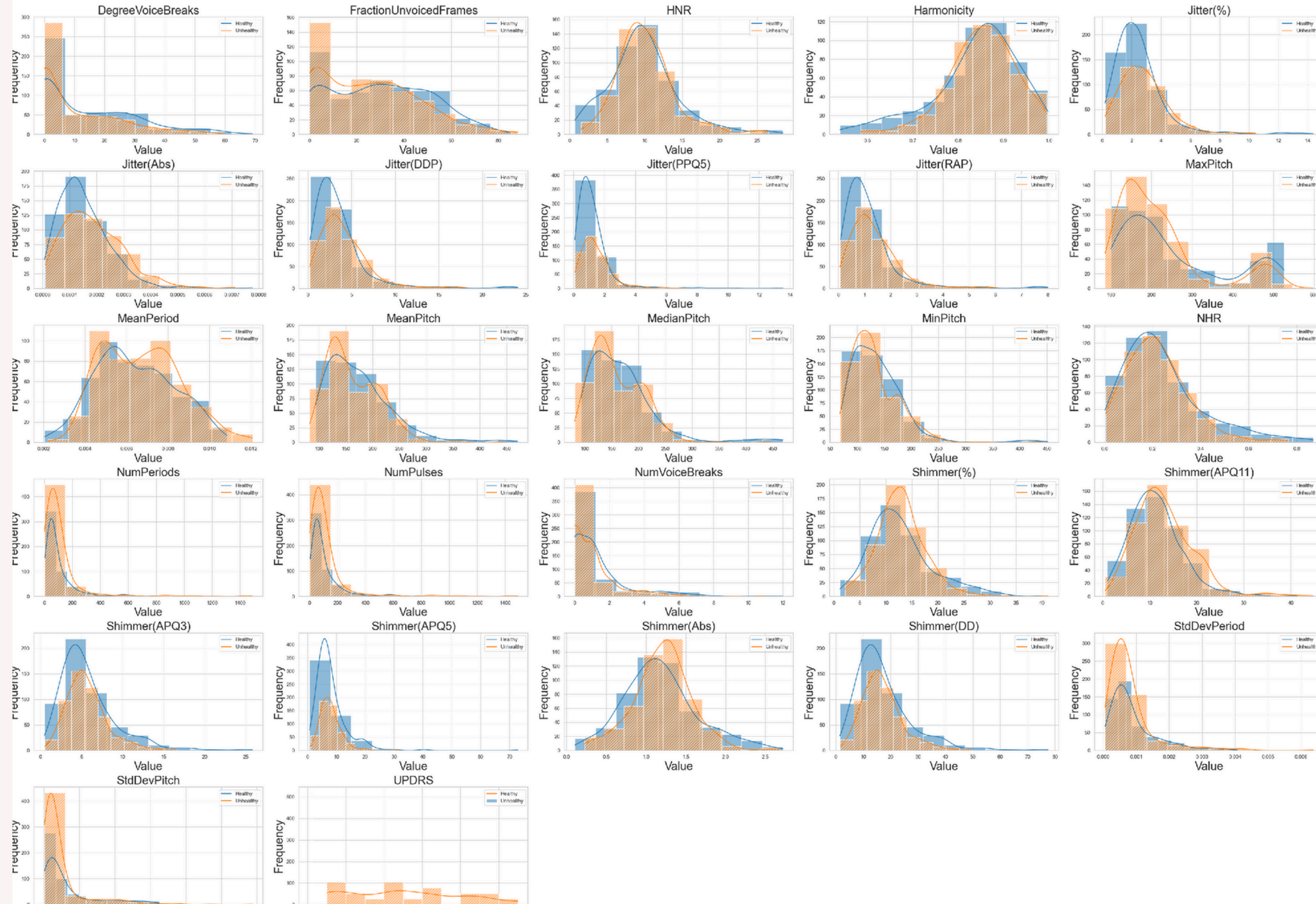
Dataset Description

It features 1039 entries and 29 columns with varied acoustic elements derived from voice recordings for each entry. We excluded the Unified Parkinson's Disease Rating Scale(UPDRS)scores. We removed Subject Identifier, UPDRS,PD Indicator during Data Pre-processing.

Column No.	Measurement category	Description
1	Subject identifier	This number identifies a study subject
2	Jitter	Jitter in %
3	Jitter	Absolute jitter in microseconds
4	Jitter	Jitter as relative amplitude perturbation (r.a.p.)
5	Jitter	Jitter as 5-point period perturbation quotient (p.p.q.5)
6	Jitter	Jitter as average absolute difference of differences between jitter cycles (d.d.p.)
7	Shimmer	Shimmer in %
8	Shimmer	Absolute shimmer in decibels (dB)
9	Shimmer	Shimmer as 3-point amplitude perturbation quotient (a.p.q.3)
10	Shimmer	Shimmer as 5-point amplitude perturbation quotient (a.p.q.5)
11	Shimmer	Shimmer as 11-point amplitude perturbation quotient (a.p.q.11)
12	Shimmer	Shimmer as average absolute differences between consecutive differences between the amplitudes of shimmer cycles (d.d.a.)
13	Harmonicity	Autocorrelation between NHR and HNR
14	Harmonicity	Noise-to-Harmonic ratio (NHR)
15	Harmonicity	Harmonic-to-Noise ratio (HNR)
16	Pitch	Median pitch
17	Pitch	Mean pitch
18	Pitch	Standard deviation of pitch
19	Pitch	Minimum pitch
20	Pitch	Maximum pitch
21	Pulse	Number of pulses
22	Pulse	Number of periods
23	Pulse	Mean period
24	Pulse	Standard deviation of period
25	Voice	Fraction of unvoiced frames
26	Voice	Number of voice breaks
27	Voice	Degree of voice breaks
28	UPDRS	The Unified Parkinson's Disease Rating Scale (UPDRS) score that is assigned to the subject by a physician via a medical examination to determine the severity and progression of Parkinson's disease.
29	PD indicator	Value "1" indicates a subject suffering from PD. Value "0" indicates a healthy subject.

Descriptive Analysis

Histogram Comparison for Healthy vs Unhealthy



Descriptive Analysis

Overview 01

the properties associated with jitter measurements, such as 'Jitter(%)', 'Jitter(Abs)', and 'Jitter(RAP)', have positively skewed distributions, indicating longer tails toward higher values

Overview 02

The 'Shimmer' features, which measure voice variability, show varied degrees of spread and skewness, particularly 'Shimmer(APQ5)'.

Overview 03

'Harmonicity' has a mean of around 0.85 and a standard deviation of 0.09, indicating a pretty consistent distribution. The 'NHR' (noise-to-harmonics ratio) has a positively skewed distribution, indicating that data skews toward higher values

Overview 04

Features such as 'NumPulses' and 'NumPeriods' have a strong positive skew, indicating potential outliers or variability. Meanwhile, metrics like 'MeanPeriod' and 'StdDevPeriod' have distributions centered near zero, indicating limited variance.

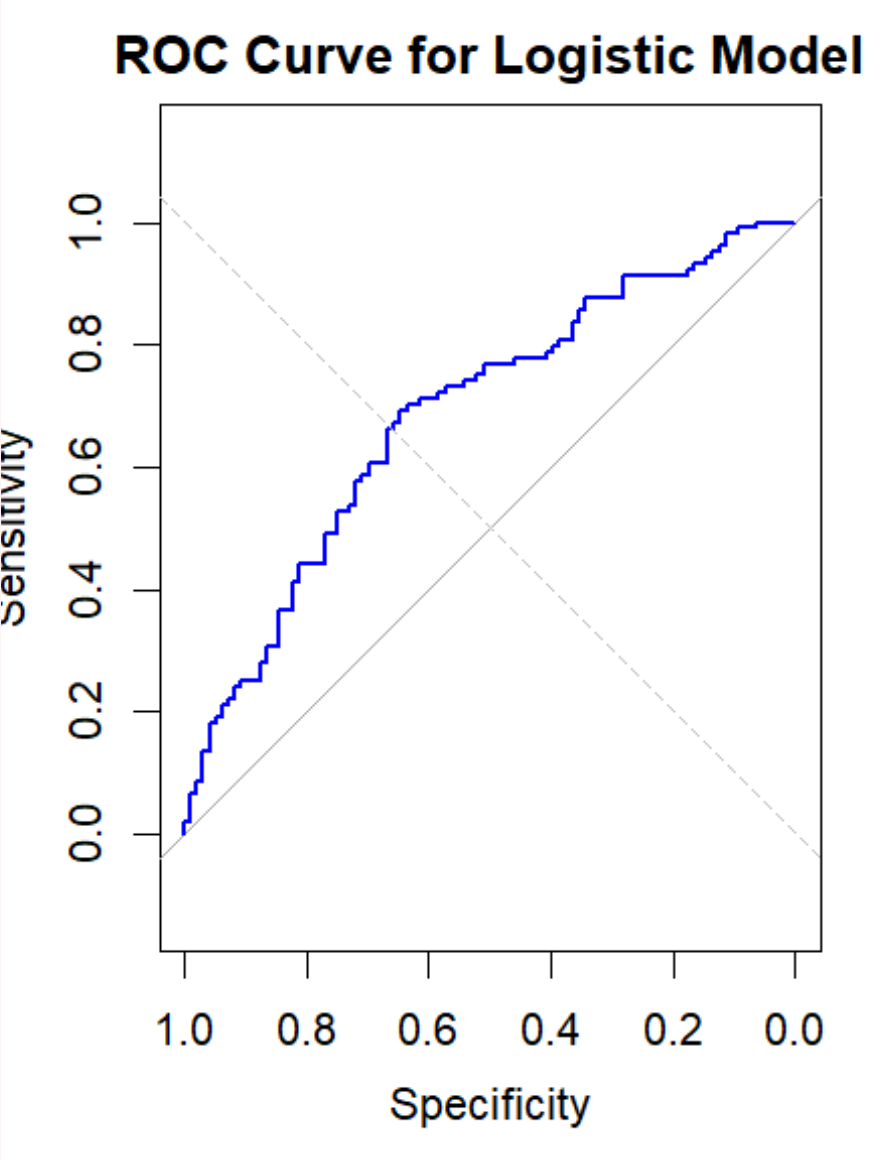
MODEL EVALUATION

LOGISTIC REGRESSION

Results

Accuracy: 0.665
Sensitivity: 0.7115
Specificity: 0.6146
F1 Score: 0.6884
ROC AUC: 0.6852

TARGET \ OUTPUT	Class0	Class1
	Class0	Class1
Class0	59 29.50%	30 15.00%
Class1	37 18.50%	74 37.00%



RANDOM FOREST

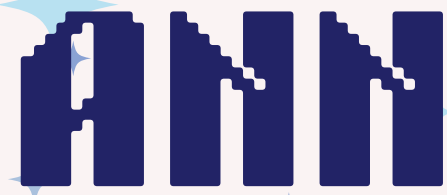
Results

Accuracy: 0.71
Sensitivity: 0.633
Specificity: 0.782
F1 Score: 0.690
ROC AUC: 0.708

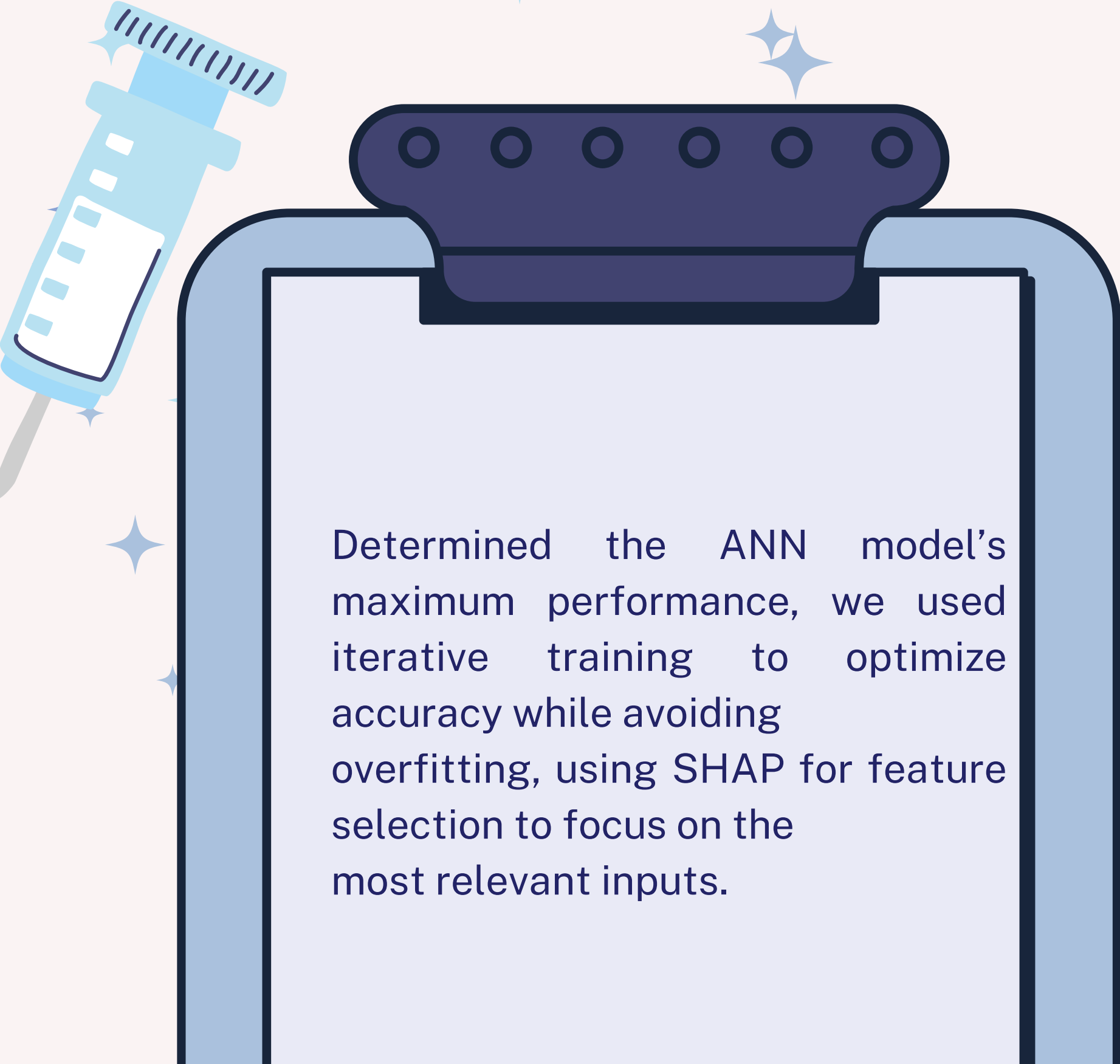
Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-2.719e+01	7.934e+00	-3.427	0.000610	***
jitter_.local.absolute.	9.149e+03	1.240e+03	7.378	1.6e-13	***
shimmer_.local.dB.	9.625e-01	3.731e-01	2.580	0.009882	**
shimmer_.apq11.	6.560e-02	1.969e-02	3.332	0.000864	***
AC	2.785e+01	7.959e+00	3.500	0.000466	***
NTH	1.018e+01	4.117e+00	2.474	0.013358	*
max_pitch	-2.552e-03	9.528e-04	-2.679	0.007387	**
no_of_periods	1.458e-03	6.421e-04	2.270	0.023215	*
mean_period	-1.773e+02	5.945e+01	-2.982	0.002861	**
frac_locally_unvoiced_frames	-1.244e-02	5.183e-03	-2.401	0.016359	*
degree_of_voice_breaks	-1.337e-02	6.676e-03	-2.003	0.045138	*

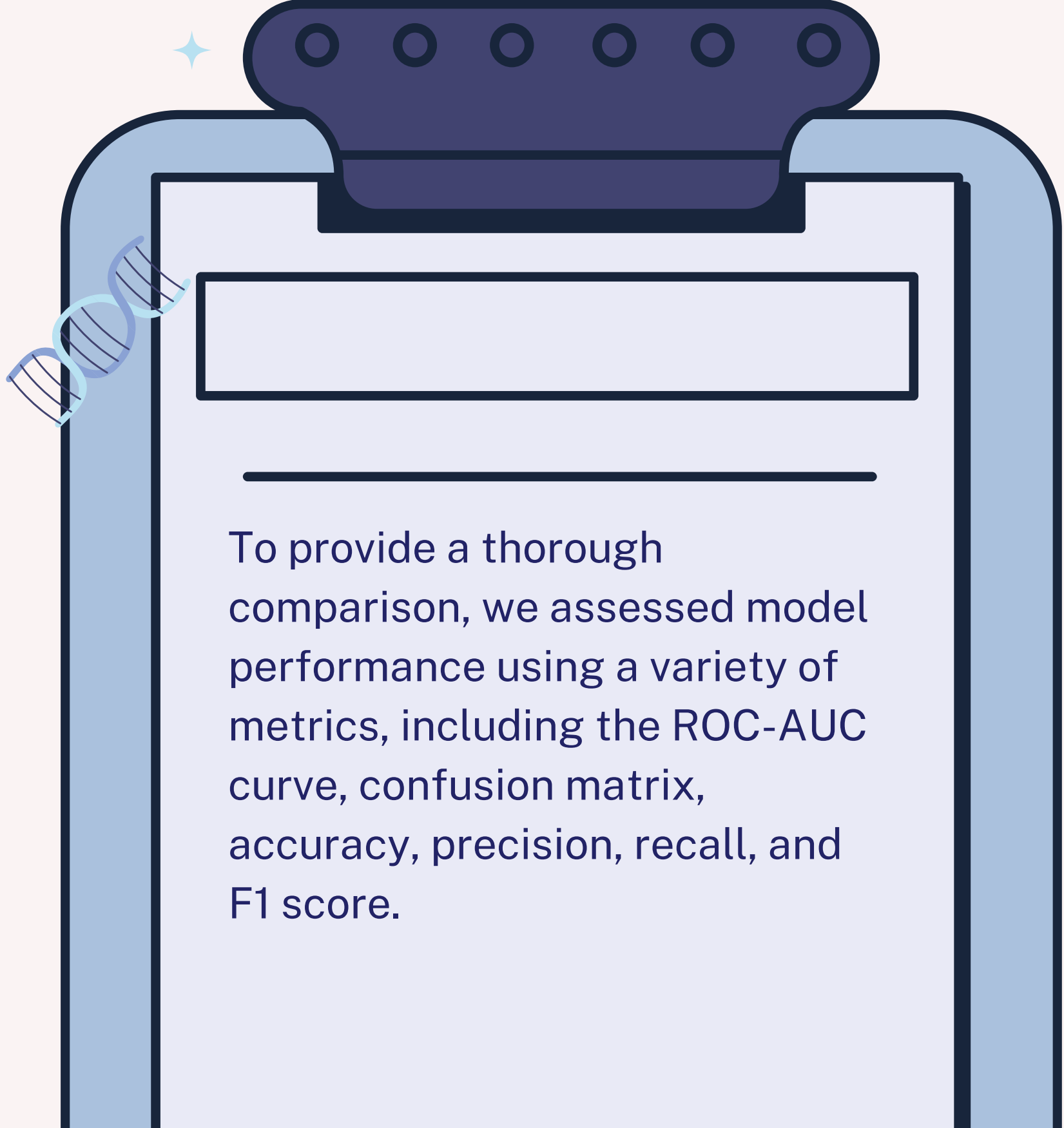
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					



ANN



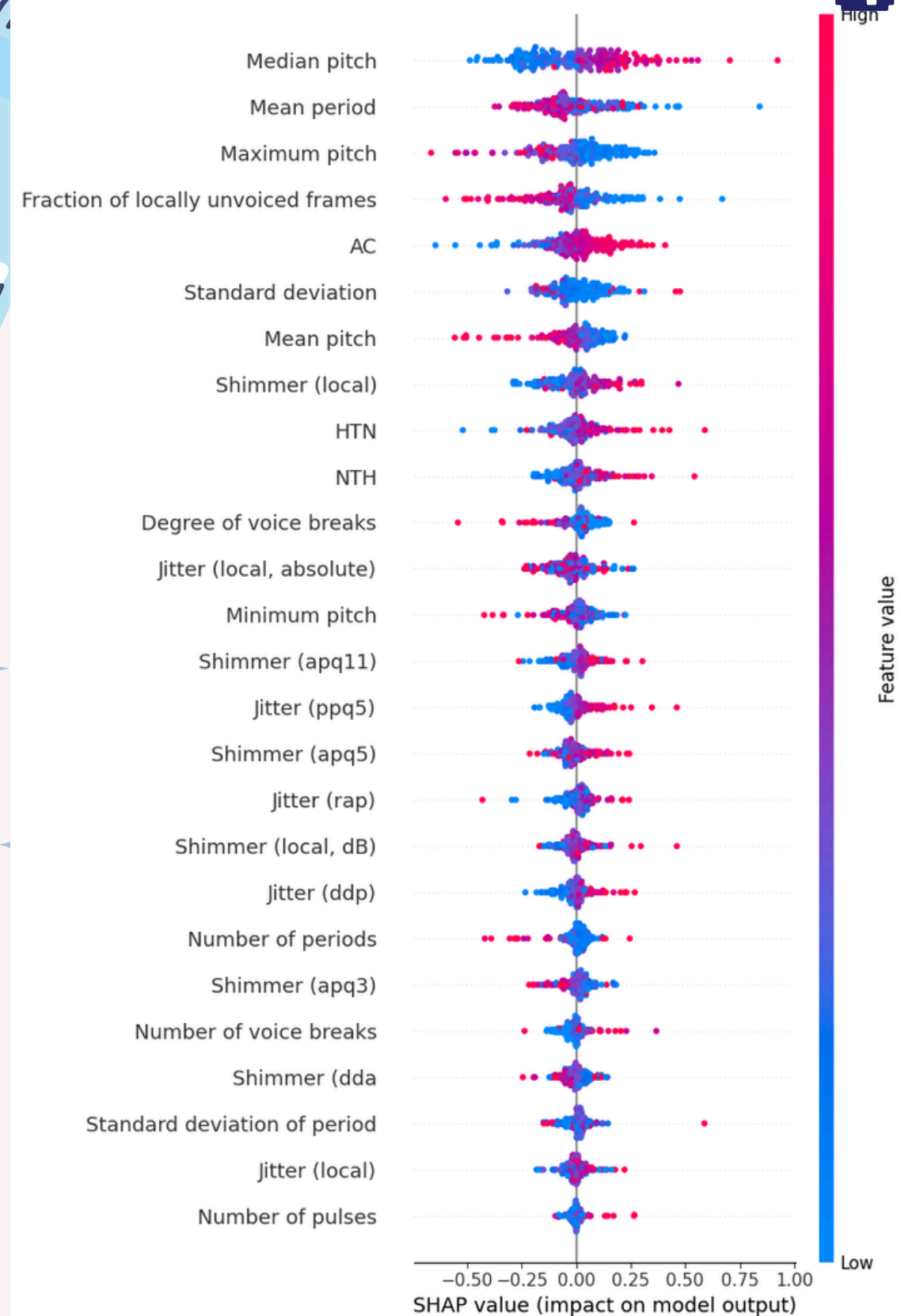
Determined the ANN model's maximum performance, we used iterative training to optimize accuracy while avoiding overfitting, using SHAP for feature selection to focus on the most relevant inputs.



To provide a thorough comparison, we assessed model performance using a variety of metrics, including the ROC-AUC curve, confusion matrix, accuracy, precision, recall, and F1 score.

MODEL EVALUATION

SHAP Analysis



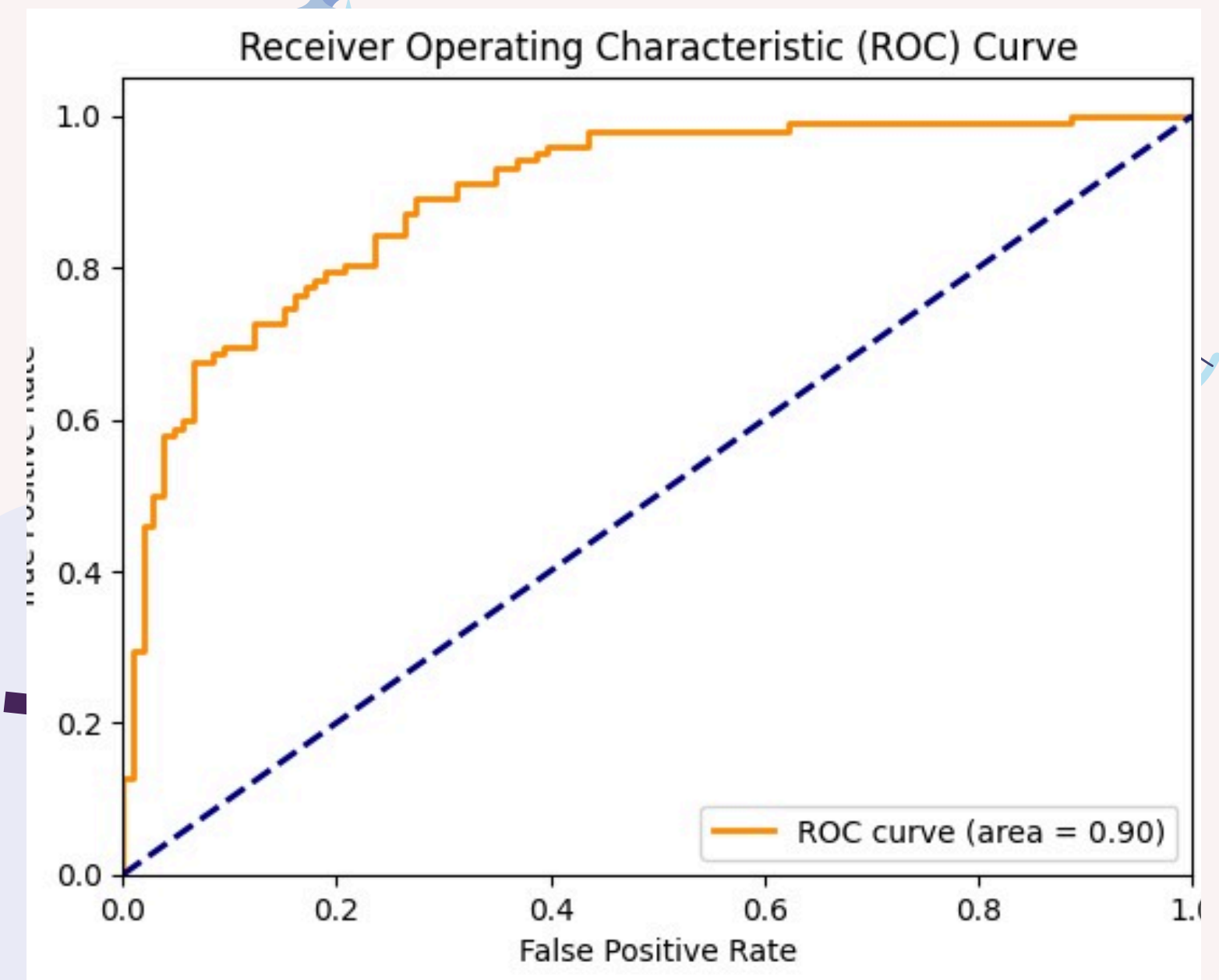
- Features like "Median pitch," "Maximum pitch," and "Mean period" have a wider spread in SHAP values, indicating significant influence on model output.
- Features with tighter distributions around zero have a smaller impact on the model's predictions.

RESULTS

The ANN model had a high accuracy of 79.83%, a sensitivity of 80.39%, and a specificity of 78.30%, showing a balanced capacity to properly identify both Parkinson's affected and unaffected patients

The AUC-ROC score was 89.63%, indicating the model's high discriminative potential

These findings highlight the model's effectiveness and reliability in detecting Parkinson's disease, indicating its potential use in diagnostic situations.



**Thank you for
your attention**

