01

Leveraging Voice for Early Detection of Chronic Kidney Disease: Enabling Continuous Monitoring in Remote Healthcare

Kangbeen Ko
Gwangju Inst. of Science and Technology
Gwangju, Republic of Korea
eyeoftyphoon@gm.gist.ac.kr
0009-0006-5286-9842

Jiwon Ryu
Seoul National University Bundang Hosp.
Seongnam, Republic of Korea
bboddo5@hanmail.net
0000-0002-2372-8948

Sejoong Kim
Seoul National University Bundang Hosp.
Seongnam, Republic of Korea
sejoong2@snu.ac.kr
0000-0002-7238-9962

02

Problem

- **The kidney disease** is a state that **decreasing of the glomerular filtration rate(GFR)**, which refers to **the degree of waste removal in the kidney** is defected, or **structural or functional decrease** of the kidney are shown.
- A case where this decrease in kidney function lasts for **more than 3 months** is called the **Chronic Kidney Disease (CKD)**, ultimately causing the need for **dialysis** and **renal replacement therapy.**
- Since decreased renal function is **not easily recognized**, **regular checkups** and **continuous monitoring** are important.
- However, existing diagnostic methods, relying solely on **blood and urine tests**, pose challenges for **continuous monitoring and diagnosis**.
- Accordingly, **an auxiliary indicator** capable of diagnosing CKD **non-invasive** and **repeatedly** is needed.

03

Hypothesis

- Based on clinical studies (Kumar et al., 2010 & Zaky et al., 2020) that **decreased renal function could affect lung and vocal cord function**, we were able to hypothesize that **the patient's voice could be an auxiliary indicator**.
- If a patient's **health indicator that can be measured without a hospital visit** is added, its diagnosing performance will increase without clinical measurements.

Table 1. Mean and SD of Acoustic Parameters for Vowel /a/ in Both the Groups (Males)

Parameters	Normal Group		Clinical Group	
	Mean	SD	Mean	SD
Fundamental frequency (F_0)	127.132	5.984	148.232	4.326
Jitter factor	0.797	0.297	2.534	0.487
Shimmer	0.6226	0.422	2.321	0.567
Rate of fluctuation in F_0	3.547	0.225	7.812	1.154
Extent of fluctuation in F_0	3.594	0.135	6.345	0.623
Rate of fluctuation in amplitude	0.730	0.213	4.245	0.232
Extent of fluctuation in amplitude	1.305	0.412	4.645	0.234

Kumar, R. B., & Bhat, J. S. (2010). Voice in chronic renal failure. Journal of Voice, 24(6), 690-693.

04

Data Descriptions

- Participants recorded six sentences, including a self-introduction.
  - 538 individuals aged 20+ included in the study.
  - Dataset contains 887 records of hospital visits, including initial and follow-up appointments.
  - Follow-ups were scheduled quarterly

TABLE I
DATASET DESCRIPTION WITH MERGED CELLS.

Stages	# of patients	# of visits	Severity
CKD 0	52	70	Non-critical
CKD 1	77	104	
CKD 2	135	205	
CKD 3	178	339	Critical
CKD 4	75	141	
CKD 5	20	28	

05

Data Description: Health

- Select health indicators that can be known **without visiting the hospital**
- SHAP assumes **independence** between features, so choose only BMI out of weight, height, and BMI
- Indicators related to vital sign, such as SBP, DBP, and PR, are **more difficult to measure** and **fluctuate** than other indicators. Therefore, it is excluded because it is **not suitable for the purpose** of easy remote monitoring

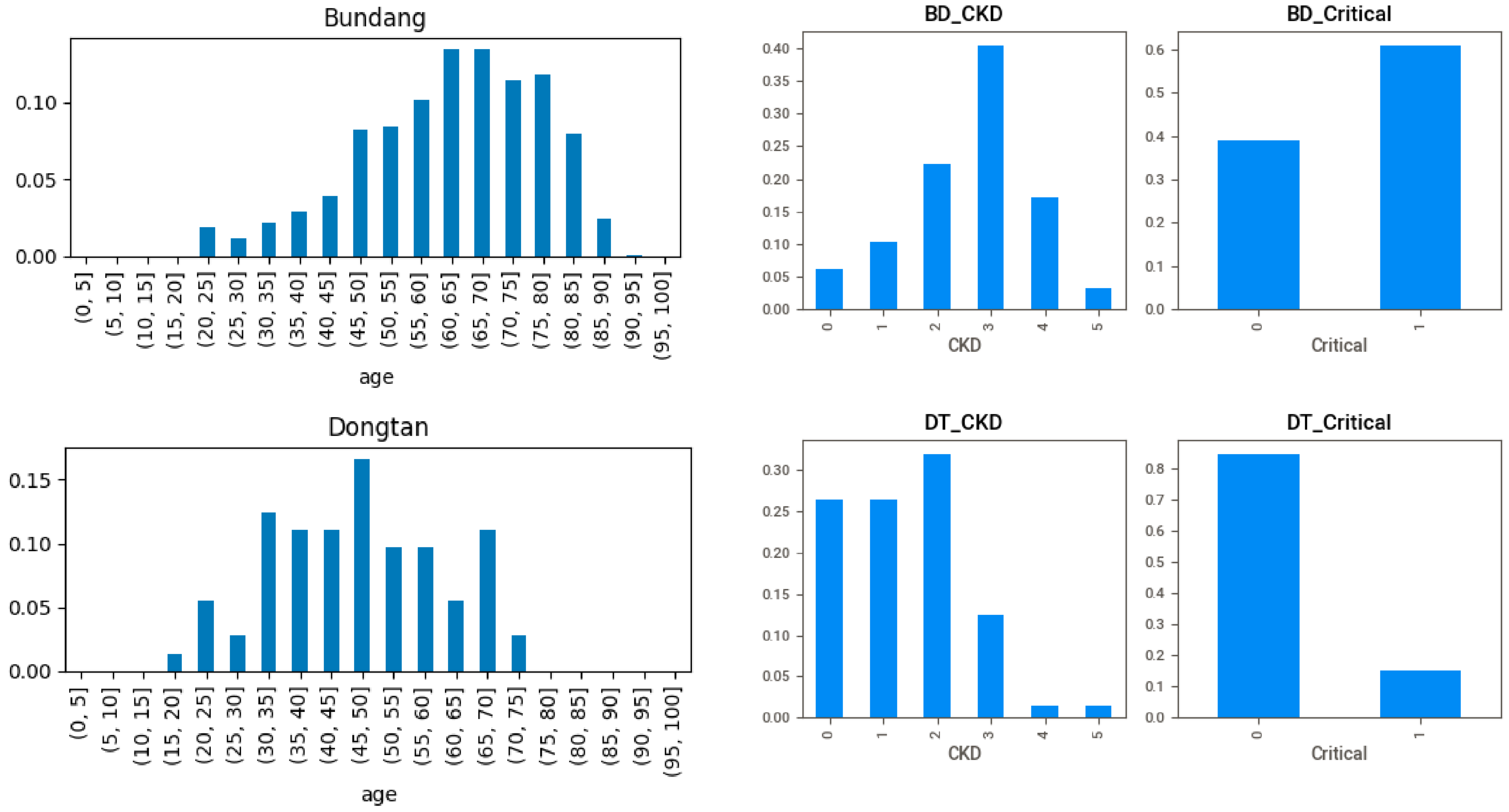
sex	age	BMI
int	int	float

Hyper-tension	Diabetes	Heart Failure	Cancer	CVD	CVA
bool (True: 1 False: 2)	bool	bool	bool	bool	bool

06

Simple EDA

- Patient data collected from two hospitals (Bundang, Dongtan) were used
- CKD 0~2 is Non-critical(0) and 3~5 is Critical(1)
- The distribution of each group is quite different, and the number of samples from Dongtan is significantly less than the data from Bundang



Experiments Setting

- **Three major experiments** were conducted as follows to prove that speech can be used as an auxiliary indicator in Critical judgment.
- All experiments were conducted for **all the data settings** mentioned in the following slides.
- **SVC, XGBoost, Liner Regression** and **TabNet** are used and **the WCE Loss** was implemented to deal with class imbalance
- **SHAP** and **PDP** is used for calculate feature importance and interpret ML Models

Step 1:

Check whether it is possible
to determine CKD severity
with simple health indicators
without blood and urine tests

Step 2:

Evaluate the possibility of
diagnosing CKD with voice only

Step 3:

See performance improvements when
using
both voice and health information

Correlation Analysis

- Conducted preliminary analysis to determine correlation between CKD severity (binary variable) and voice features (numerical variables).
- Calculated point-binary correlation coefficient and p-value.
- Identified 72 significant features with p-value < 0.05.
- Top 10 features include: Age, Hypertension, Diabetes, MeanUnvoicedSegmentLength, spectralFlux sma3 stddevNorm, etc.

$$r_{pb} = \frac{M_1 - M_0}{s_n}\sqrt{\frac{n_1 n_0}{n^2}}$$

TABLE II
TOP-10 FEATURES CORRELATED WITH TARGET VARIABLE.

Rank	Feature	abs(r_{pb})	p-value
1	Age	0.372	0.00
2	Hypertension	0.294	0.00
3	Diabetes	0.282	0.00
4	MeanUnvoicedSegmentLength	0.196	0.00
5	spectralFlux sma3 stddevNorm	0.192	0.00
6	MFCC3 sma3 amean	0.166	0.00
7	loudness sma3 stddevNorm	0.165	0.00
8	MFCC3V sma3nz amean	0.162	0.00
9	loudness sma3 percentile20.0	0.158	0.00
10	VoicedSegmentsPerSec	0.157	0.00

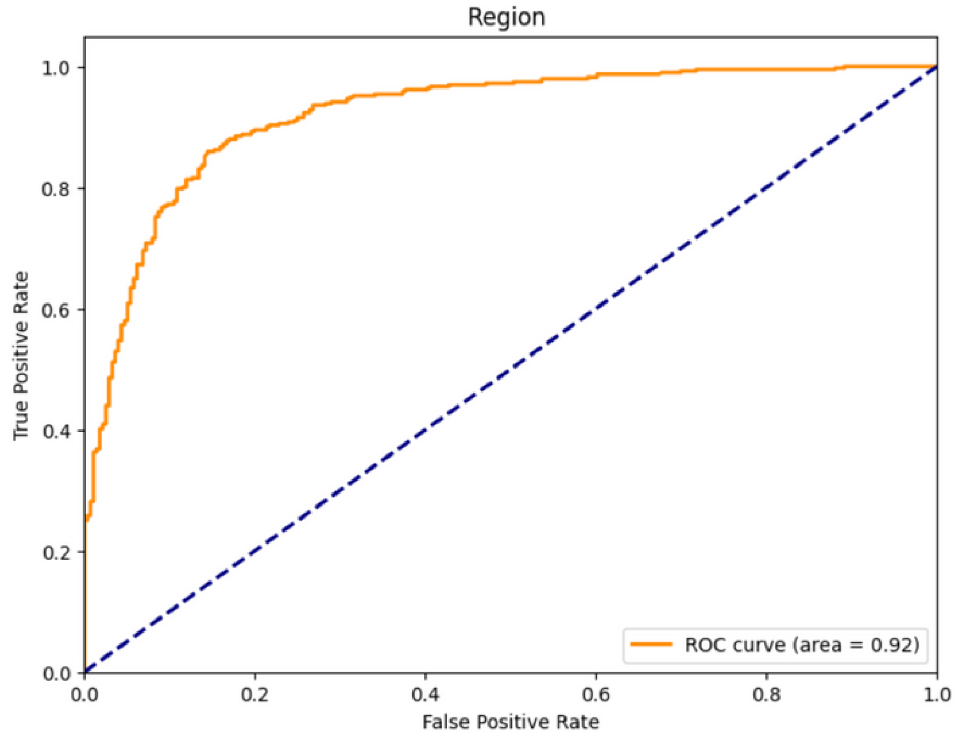
09

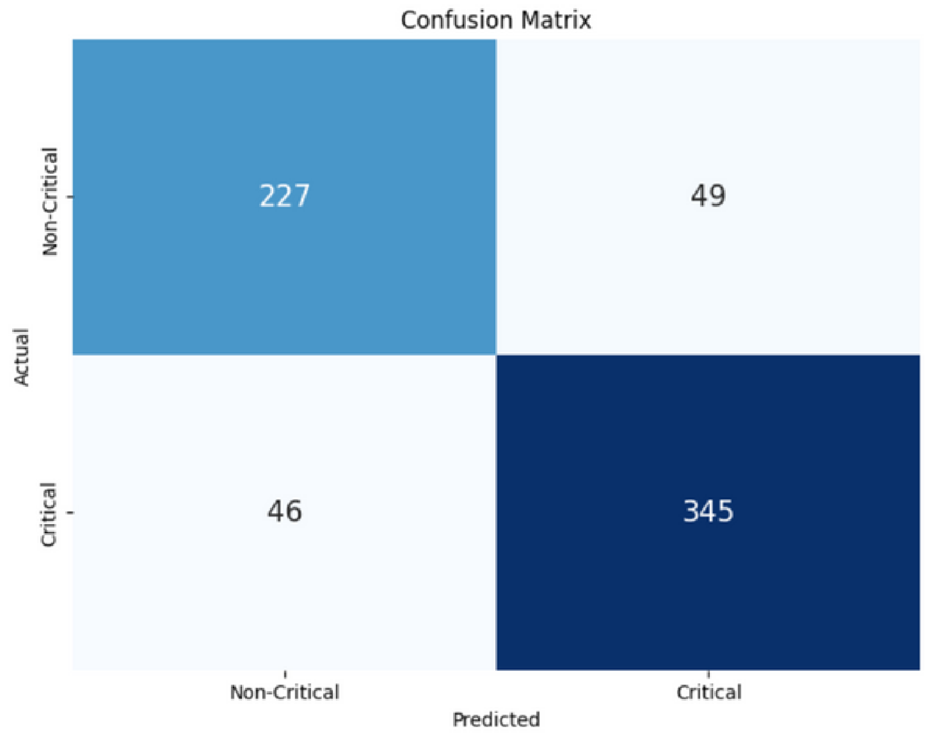
Performance Evaluation

- Comparative experiments using health indicators (age, HTN, DM) and vocal features for CKD severity classification.
- **Health I (age, HTN, DM)**
- **Health II (expanded health indicators)**
- **Superior performance when combining vocal features with health information.**

TABLE III
PERFORMANCE EVALUATION.

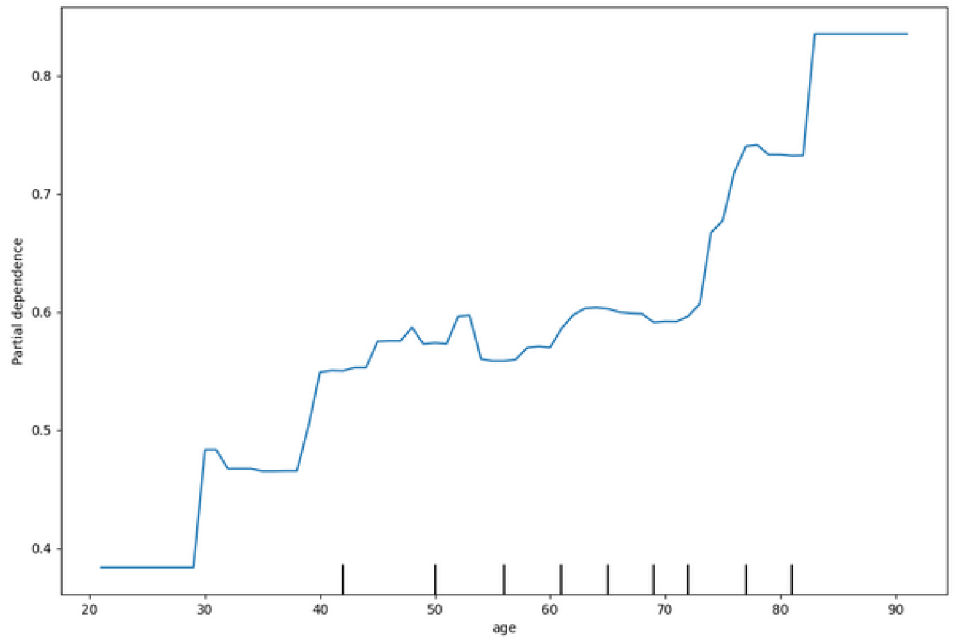
Feature Set	Model	Precision	Recall	F1	AUROC
Voice Only	SVC	0.719	0.68	0.698	0.73
Voice Only	XGBoost	0.757	0.731	0.706	0.76
Health I	SVC	0.74	0.686	0.71	0.73
Health I	XGBoost	0.795	0.754	0.743	0.84
Health I & Voice	SVC	0.854	0.777	0.814	0.88
Health I & Voice	XGBoost	**0.876**	0.816	0.826	0.90
Health II & Voice	SVC	0.835	0.83	0.811	0.91
Health II & Voice	XGBoost	**0.876**	**0.882**	**0.857**	**0.92**





Explainable AI

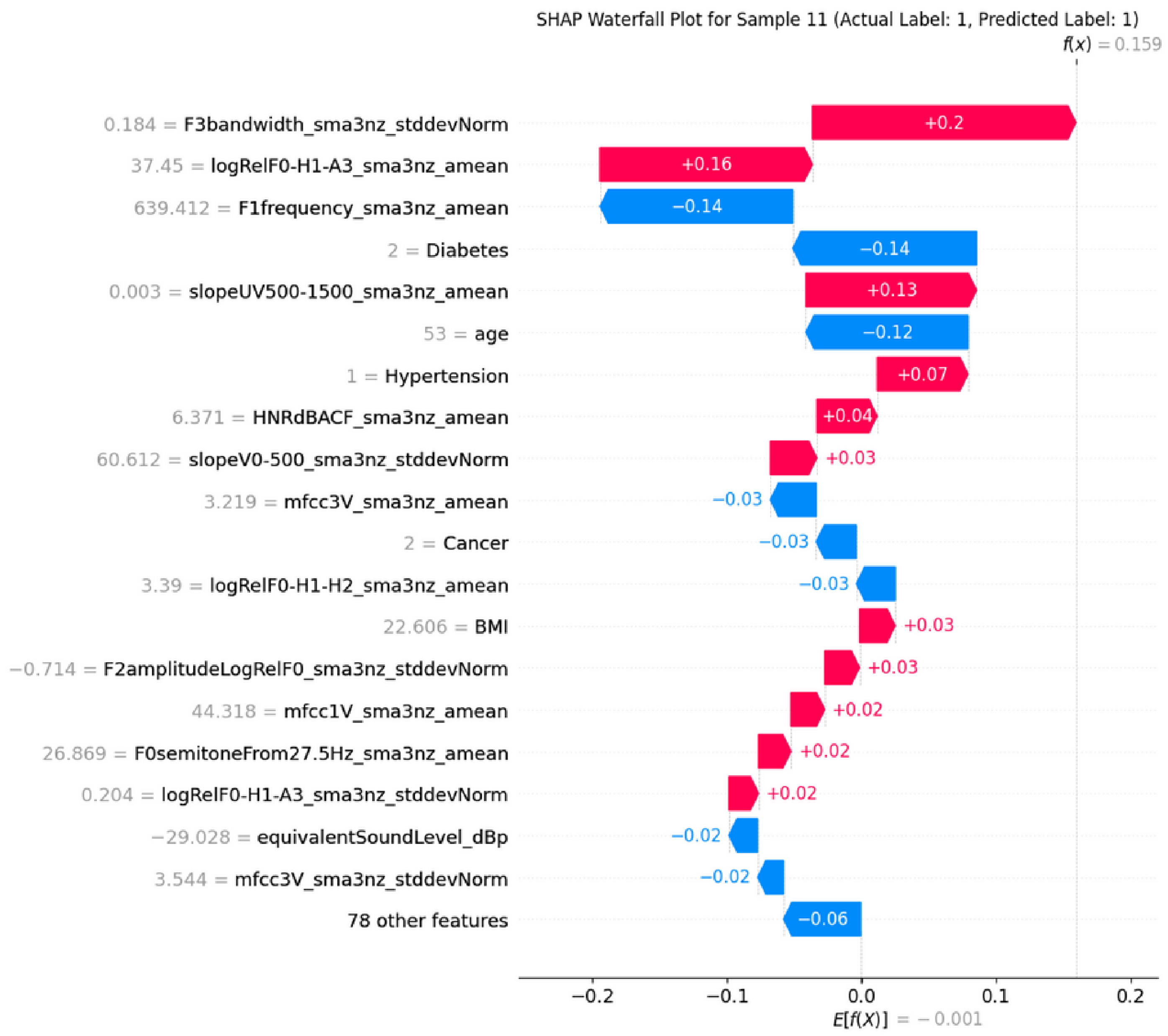
- Utilized Partial Dependence Plots (PDPs) and SHAP values to evaluate feature significance and impact on predictions.
- Combined PDP and SHAP analyses enhanced model interpretability and confirmed its clinical utility.
- Example analyses demonstrated how voice features improved prediction accuracy, even when basic health indicators were ambiguous.
-

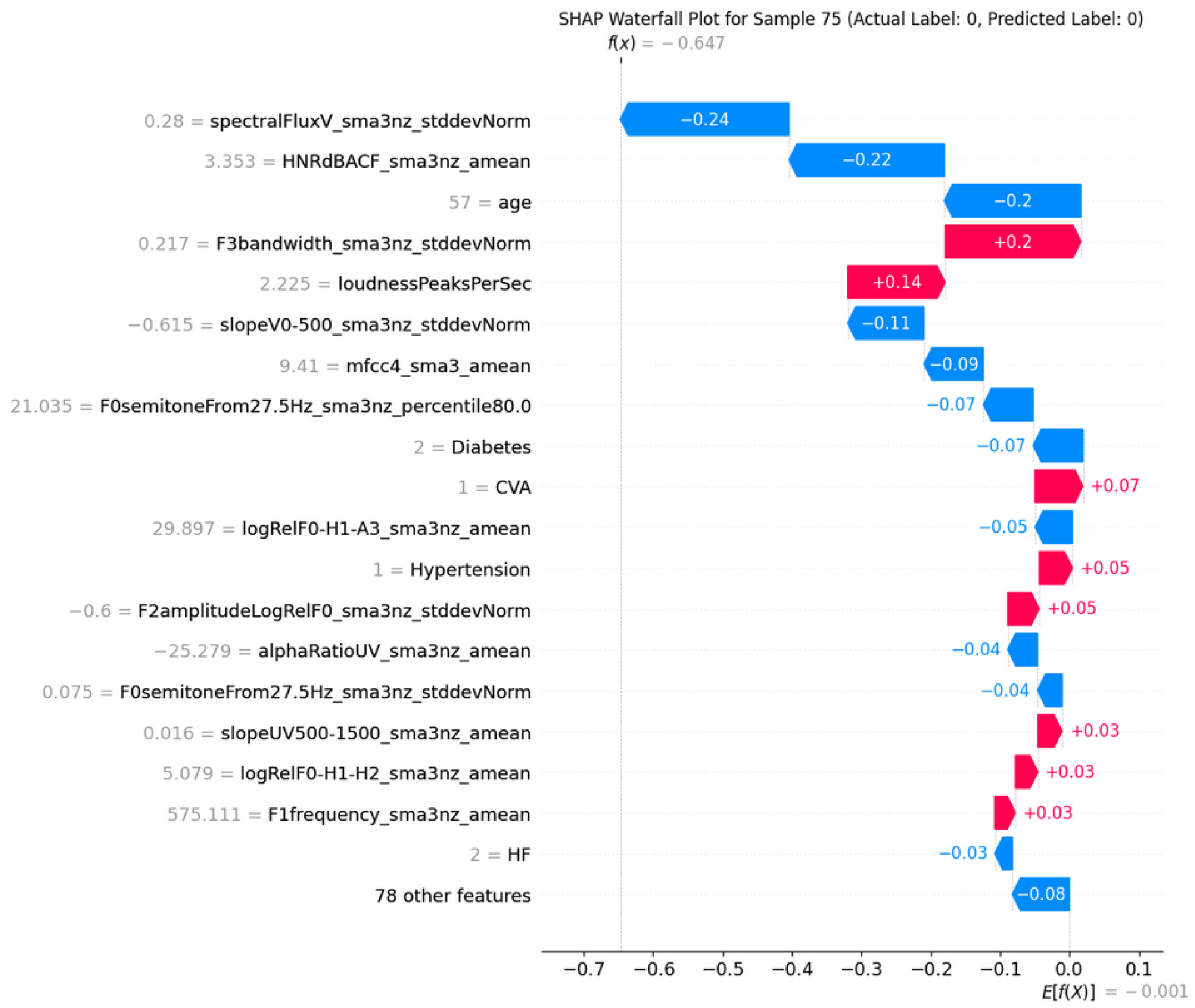


PDP for Ages

SHAP values

SHAP Examples for Individual Predictions





Conclusions & Future Works

- Explored patient voice features as biomarkers for CKD diagnosis alongside traditional methods.
- Demonstrated the accuracy of voice features in classifying CKD severity using machine learning.
- Potential for continuous and remote patient monitoring, particularly for severe cases.
- Plans to refine classification by detailing CKD severity stages, using advanced machine learning and deep learning models.
- Consideration of spectrogram-based voice imaging techniques with CNNs for precise analysis.
- Acknowledgment of the dataset limitation and the need to expand and augment data for more comprehensive research.

Thank you

If You have Any Questions, Feel Free to Ask.

e-mail: eyeoftyphoon@gm.gist.ac.kr