**Voice Authentication System**

**Introduction**

This document details the development of a real-time voice authentication system that leverages Fast Fourier Transform (FFT) and machine learning. The system is designed to enhance security by verifying users based on their voice characteristics. It has applications in access control, banking security, IoT devices, and smart assistants.

**Technical Architecture**

**Workflow Overview**

1. **Audio Capture** – The system records a short audio sample.
2. **Preprocessing** – Noise reduction, voice activity detection (VAD), and feature extraction.
3. **Feature Extraction** – Extracting Mel-frequency cepstral coefficients (MFCCs), chroma features, spectral contrast, and tonnetz.
4. **Model Training** – Using an SVM classifier trained with voice samples.
5. **Authentication** – Comparing real-time input with trained models to determine authenticity.

**Fast Fourier Transform (FFT) Usage**

FFT is applied to convert time-domain audio signals into frequency-domain features, aiding in better feature extraction and noise filtering.

**Implementation Details**

**Key Algorithms & Libraries**

* **pyaudio** – Captures real-time audio.
* **wave** – Handles audio file operations.
* **librosa** – Performs signal processing and feature extraction.
* **scikit-learn** – Machine learning framework for training SVM models.
* **joblib** – Saves and loads trained models.

**Dependencies**

* Python 3.x
* Required libraries: numpy, pyaudio, wave, joblib, librosa, scikit-learn

**Dataset & Training**

* **Dataset Path:** C:\Users\atish\Documents\VoiceAuth\Audios\user\_0
* **Preprocessing Steps:**
  + Noise reduction using low-pass filtering.
  + Trimming silence using librosa’s VAD.
  + Feature extraction using MFCCs and additional features.
* **Training Methodology:**
  + User and impostor voice samples are collected.
  + Features are extracted and scaled using StandardScaler.
  + Hyperparameter tuning using GridSearchCV.
  + Support Vector Machine (SVM) model is trained and stored.

**Real-Time Authentication Process**

1. User records a fresh voice sample.
2. The system extracts features and normalizes them.
3. Features are passed through the trained SVM model.
4. Decision based on probability thresholds (e.g., confidence >90%).
5. Authentication result (Access Granted/Denied) is displayed.

**Performance Metrics**

* **Accuracy:** Measured through test datasets.
* **False Acceptance Rate (FAR):** Probability of impostors being accepted.
* **False Rejection Rate (FRR):** Probability of genuine users being denied.
* **Latency:** Time taken for authentication processing.

**Comparison with Existing Systems**

| **Feature** | **This System** | **Traditional Voice Authentication** |
| --- | --- | --- |
| FFT-Based Feature Extraction | Yes | Limited |
| Machine Learning Integration | Yes | Basic Algorithms |
| Custom Dataset Training | Yes | Predefined |
| Real-Time Processing | Yes | Some delay |

**Security & Vulnerabilities**

**Risks & Mitigation Strategies**

* **Spoofing Attacks** – Implementing additional anti-spoofing measures.
* **Noise Interference** – Using adaptive noise filtering.
* **Model Overfitting** – Expanding dataset diversity.

**Installation & Usage Guide**

1. Install dependencies using pip install numpy pyaudio wave joblib librosa scikit-learn.
2. Place dataset in C:\Users\atish\Documents\VoiceAuth\Audios.
3. Run python voice\_auth.py to train the model.
4. Use authenticate\_voice() function to verify users.

**Use Cases & Applications**

* **Access Control** – Securing entry points.
* **Banking Authentication** – Secure voice-based transactions.
* **IoT & Smart Devices** – Personalized voice interactions.

**Scalability & Deployment**

* **Cloud Integration:** Train and deploy on cloud servers.
* **Embedded System Adaptation:** Optimize model for edge devices.
* **Scalability:** Can handle larger datasets with distributed training.

**Challenges & Future Improvements**

* **Enhancing Accuracy:** Adding deep learning-based models.
* **Reducing Latency:** Optimizing feature extraction pipeline.
* **Better Security Measures:** Implementing advanced spoof detection.

**Conclusion**

This system provides a robust and scalable voice authentication solution with real-time processing capabilities. Future enhancements can include deep learning integration and improved noise-handling techniques.

**Appendix & References**

* [Librosa Documentation](https://librosa.org/)
* [Scikit-learn SVM Guide](https://scikit-learn.org/stable/modules/svm.html)