**Tasks:**

**Task 1: Find and fix any bugs within base.py, database.py, main.py, schemas.py**

|  |  |  |  |
| --- | --- | --- | --- |
| **File** | **Bugs** | **Fix Implemented** | **Code Fix** |
| **base.py** | No bugs found | No changes done |  |
| **database.py** | Typo in AudioMetadat import. | Imported correct class name and moved it into import section | from sqlalchemy import create\_engine from sqlalchemy.orm import sessionmaker from app.base import Base from app.models import AudioMetadata |
| **main.py** | Enigne import is not used anywhere | Removed engine import | from app.database import SessionLocal # engine - removed as it is not used |
| Length of audio data can be float, so integer division will not be precise | Changed from integer division to fractional division | length\_seconds = len(audio\_data) / SAMPLE\_RATE |
| **models.py** | Timestamp column is defined as a String, but it should be a DateTime type | Changed timestamp column to DateTime and pulled all the required imports | from sqlalchemy import Column, Integer, String, Float, DateTime from app.base import Base from datetime import datetime, timezone   timestamp = Column(DateTime, default=lambda: datetime.now(timezone.utc)) |
| length\_seconds column is defined as integer, but it should be a Float | Changed length\_seconds column to float for precise results | length\_seconds = Column(Float) |
| **schemas.py** | Field import as it is not used anywhere in the code | Removed Field from the import as it is not used | from pydantic import BaseModel, field\_validator |
| Formula used to find audio duration is seen wrong and it is not consistent with main.py file | Formula for audio duration should be len(audio\_data)/sample rate) same as main.py | min\_length\_samples = SAMPLE\_RATE  if len(audio\_data) < min\_length\_samples:  raise ValueError(f"Audio file is too short. Minimum length is {len(audio\_data) / SAMPLE\_RATE} seconds.") |
| **All files** | Many lint errors observed | All the necessary lint errors are fixed |  |

Please find attached all the bug fixed files in the email.

**Task 2:** Write a brief report if any bugs are found, priorities up to four bugs in order of most critical, explaining the rationale for each.

**Bug Report based on priority**

**Summary**

The code files provided (base.py, database.py, main.py, models.py, and schemas.py) were analyzed for bugs and inconsistencies. Below is a prioritized list of the most critical issues identified, along with their rationale and fixes.

1. **Inconsistent timestamp Type in models.py**

* **Bug**: The timestamp column in AudioMetadata is defined as a String. This is inappropriate for handling date and time values.
* **Rationale**: Using String prevents proper date-time operations like sorting or filtering.

This can cause incorrect behavior in database queries.

This impacts database functionality, preventing proper date-time operations. Critical for ensuring application correctness.

* **Fix**: Changed the column to DateTime with a default value of datetime.now().
* **Priority**: Critical
* **Severity**: Critical

1. **Integer Division in Audio Length Calculation in main.py**

* **Bug**: The length of audio files is calculated using integer division (//). This truncates fractional seconds, leading to imprecise results.
* **Rationale**: Causes incorrect audio duration calculations, which can affect audio processing logic especially for smaller files. Needs immediate attention for accurate results.
* **Fix**: Changed integer division to floating-point division (/) for accurate audio length calculation.
* **Priority**: High
* **Severity**: High

1. **Inconsistent Formula for Audio Length Validation in schemas.py**

* **Bug**: The formula for validating audio length is inconsistent between schemas.py and main.py.

The schema uses a hardcoded value (4000) instead of referencing SAMPLE\_RATE.

* **Impact**: Incorrect formula leads to Validation mismatch and inaccurate results.

And any change to SAMPLE\_RATE in main.py would not reflect in schemas.py, causing potential validation mismatches.

Less critical but important for maintaining consistency and inaccuracies.

* **Fix**: Updated the formula in schemas.py and used SAMPLE\_RATE from main.py instead of hardcoding it
* **Priority**: High
* **Severity**: High

1. **Typo in database.py Import (AudioMetadat)**

* **Bug**: The AudioMetadat class is imported incorrectly due to a typo. The correct class name is AudioMetadata.
* **Impact**: This results in an ImportError, breaking the functionality of table creation and halting the system during initialization. This is significant but easier to fix compared to the above issues.
* **Fix**: Fixed the typo in the import
* **Priority**: High
* **Severity**: critical

**Task 3**: Create Unit Tests in any Python testing framework of your choice.

Please find attached the unit test case files test\_databse.py, test\_schema.py, test\_endpoints.py in the email.

