**Eötvös Loránd Tudományegyetem**

**Informatikai Kar**

**Informatikatudományi Intézet**

**Programozási Nyelvek és FordítóprogramokTanszék**

Magyar népzene felismerésére szolgáló Blueprint fejlesztése

Szerző: Témavezető:

Fegyó Benedek Dr. Tejfel Máté

Programtervező informatikus BSc. Egyetemi docens

**Budapest 2024**

Ide kerül a hivatalos témabejelentő lap.

Tartalomjegyzék

[1. Bevezetés 1](#_Toc180677579)

[2. Felhasználói dokumentáció 1](#_Toc180677580)

[2.1. Probléma rövid megfogalmazása 1](#_Toc180677581)

[2.2. Felhasznált módszerek 1](#_Toc180677582)

[2.3. Program használata 2](#_Toc180677583)

[2.3.1. Előfeltételek 2](#_Toc180677584)

[2.3.2. Virtuális Környezet telepítése teszteléshez 3](#_Toc180677585)

[2.3.3. Blueprint telepítése 4](#_Toc180677586)

[2.3.4. Zenefelismerő használata 5](#_Toc180677587)

[2.4. Futás közben adódó tipikus hibák megoldása 7](#_Toc180677588)

[3. Fejlesztői dokumentáció 8](#_Toc180677589)

[3.1. A probléma specifikálása 8](#_Toc180677590)

[3.2. MQTT 8](#_Toc180677591)

[3.3. Crepe 8](#_Toc180677592)

[3.4. Python használata a Blueprint és az alkalmazás megírásához 8](#_Toc180677593)

[3.5. Virtuális környezet (*Virtual\_ENV/…*) 8](#_Toc180677594)

[3.5.1. Multipass 8](#_Toc180677595)

[3.5.2. SSH kulcsok generálása 9](#_Toc180677596)

[3.5.3. Cloud-Init.yaml fájl létrehozása 9](#_Toc180677597)

[3.5.4. A virtuális gépek létrehozása és konfigurálása (*setup\_virtual\_env.bat*) 10](#_Toc180677598)

[3.6. Blueprint (*Blueprint/…*) 12](#_Toc180677599)

[3.6.1. Control Node konfigurálása (*control\_node\_setup.sh*) 14](#_Toc180677600)

[3.6.2. Node-ok konfigurálása (*node\_setup.sh*) 14](#_Toc180677601)

[3.7. Zenefelismerő alkalmazás (*Application*) 15](#_Toc180677602)

[3.7.1. Adatok bekérése a felhasználótól (*lampion.py*) 17](#_Toc180677603)

[3.7.2. Üzenet és fájl küldése MQTT-n Python szkripttel (*mqtt\_publish\_message.py és mqtt\_publish\_file.py*) 18](#_Toc180677604)

[3.7.3. Control Node működése az alkalmazás futása közben (*control\_node*\_*action.sh*) 19](#_Toc180677605)

[3.7.4. A felismerő node működése az alkalmazás futása közben (*node\_action.sh*) 20](#_Toc180677606)

[3.8. A felhasznált modulok és könyvtárak listája és verziószáma 26](#_Toc180677607)

[3.9. Tesztelési terv és a tesztelés eredményei 26](#_Toc180677608)

[4. Összefoglalás és további fejlesztési lehetőségek 27](#_Toc180677609)

[5. Irodalomjegyzék 28](#_Toc180677610)

[6. Melléklet 30](#_Toc180677611)

# Bevezetés

1. **Bevezetés**, mely tartalmazza

* a témaválasztás indoklását,
* és a megoldandó feladat rövid, közérthető leírását.

# Felhasználói dokumentáció

## Probléma rövid megfogalmazása

A magyar népzenében a dallamokat hallás alapján, kézről kézre adják tovább generációk óta. <> Ennek megvan a szépsége, hiszen aki szeretné és lehetősége van rá, az közvetlenül a forrástól tud tanulni és az ő értelmezésében meghallgatni akár egy egész rendet vagy dallamfűzést. Azonban annak érdekében, hogy ezek a dallamok fent tudjanak maradni olyan esetekben is, amikor valamilyen okból kifolyólag nincs, aki megtanulja vagy lekottázza az adott hanganyagot szükségessé válik egy olyan alkalmazás, amely ezt automatikusan megteszi anélkül, hogy nagyfokú emberi beavatkozást igényelne. Ezen kívül az is lényeges, hogy minimális erőforrású eszközökön is fusson. Ezeket a szempontokat azért kell figyelembe venni, mivel egy népzenével foglalkozó embertől sem az informatika beható ismerete, sem egy mesterséges intelligencia alapú felismerő futtatására elégséges személyi számítógép birtoklása nem elvárható. Ennek következtében a cél egy olyan alkalmazás kifejlesztése volt, ami automatikusan telepíti a szükséges modulokat, felépíti a hálózatot a távoli gépekkel amiken a felismerők futhatnak és egy viszonylag egyszerű és megbízható felhasználói élményt biztosít a probléma komplexitását figyelembe véve.

## Felhasznált módszerek

A probléma megoldására egy Blueprintet fejlesztettünk ki, ami egy ismeretlen, előre nem definiált környezetben, automatikusan felállítja az MQTT hálózatot a felhasználó személyi számítógépe és a felismerőket futtató virtuális gépek között, majd azokra telepíti az alkalmazásához szükséges modulokat és fájlokat. A felismerők egy tensorflow alapú music21 Python könyvtárat használnak, amelynek az eredményét feldolgozzák majd musicxml formátumúvá konvertálják és továbbítják a megfelelő Node felé.

## Program használata

A program használatának több fázisa van. Először teljesíteni kell bizonyos előfeltételeket. Ezek olyan hardveres és szoftveres kikötések, amelyeknek teljesülnie kell az alkalmazás futtatása előtt. A fizikai kikötések annak a jelenségnek a következménye, hogy bizonyos funkciók működése nagyban korlátozott, vagy lehetetlen megfelelő hardver nélkül. A szoftveres előfeltételek abból adódnak, hogy az alkalmazás célzott felhasználói operációs rendszere a Windows. Ennek következtében bizonyos modulok automatikus telepítése több komplikációt és kompatibilitási problémát okozna a jövőben, mint amennyivel megnehezíti a felhasználó dolgát.

A második fázis opcionális. Erre akkor van szükség amennyiben az alkalmazást teszt üzemmódban szeretnénk futtatni. Ebben az esetben feltelepülésre kerül egy virtuális környezet a felhasználó személyi számítógépére, amely képes arra, hogy az alkalmazást teszt üzemmódban futtassa, ideértve a különböző funkciókat betöltő virtuális gépeket. Amennyiben az alkalmazást fizikai szervereken szeretnénk futtatni, ez a fázis teljes mértékben kihagyható.

A harmadik fázis a Blueprint telepítése. Amennyiben mind a fizikai mind a szoftveres előfeltételek teljesülnek, a *nodelist.txt* helyes kitöltésével é a *Blueprint/setup.bat* batch szkript futtatásával, a Blueprint egy kattintásra feltelepül.

A negyedik fázis az alkalmazás használata. Ez a *Application/full\_lampion.bat* batch szkript futtatásával lehetséges, ami bekéri a felhasználótól a megfelelő adatokat, hangfájlokat, továbbítja őket a felismerőnek MQTT-n keresztül, majd a feldolgozott hanganyagot a Control Node egy kottává állítja össze és eltárolja.

### Előfeltételek

#### Fizikai előfeltételek

Amennyiben az alkalmazást a felhasználó a személyi számítógépén szeretné teszt üzemmódban futtatni szüksége lesz egy olyan eszközre, amely képes létrehozni a megfelelő virtuális környezetet. Mivel a felismerők TensorFlow alapúak, amihez az ajánlott minimum RAM 8 gigabájt, az ajánlott minimum háttértár pedig 10 gigabájt, így a felhasználó személyi számítógépének képesnek kell lennie ennek a felismerőnkkénti allokálására az operációs rendszer futtatásán felül. (EITCA, 2024) Amennyiben a felismerők optimális működése nem elvárt akkor ez az igény két gigabájtig csökkenthető. (3.5.4) Továbbá a Control Node és a Mosquitto Broker további 2 – 2 gigabájt RAM-ot igényel. Így a tesztüzemmódban történő futtatás minimum RAM igényét a következő képlettel lehet meghatározni.

Ellenkező esetben, ha a felhasználó személyi számítógépén nem szeretnénk virtuális felismerőket futtatni, akkor a RAM és háttértár igény megegyezik az operációs rendszer megfelelő működéséhez szükségessel.

#### Szoftveres előfeltételek

A Blueprint telepítéséhez szükséges a GitHub. Ennek az automatikus telepítése Windowson komplikálttá és kevésbé megbízhatóvá tenné a telepítési folyamatot így ennek a telepítése átkerült a szoftveres előfeltételekbe

Amennyiben az alkalmazást teszt üzemmódban szeretnénk futtatni, szükséges egy Virtuális Gép koordinátor. Ez teszi lehetővé, hogy a virtuális környezet létrehozása és használata folyékonyan történjen, ugyanakkor emuláljon egy valós környezetet, hogy a tesztelés annyira megközelítse a valóságot, amennyire lehetséges. Az alkalmazás fejlesztése és tesztelése közben a Multipass, Ubuntu alapú virtuális gép szervező volt használva.

### Virtuális Környezet telepítése teszteléshez

A virtuális környezetet, az előfeltételek teljesítése után lehet telepíteni. A *Virtual\_ENV* mappában megtalálható minden, amire ehhez szükség lesz. A mappa tartalmaz egy inicializációs konfigurációt, egy előre generált SSH kulcsot a virtuális gépek jelszó nélküli eléréséhez és a telepítő batch szkriptet. A *setup\_virtual\_environment.bat* batch fájl futtatásával felépül a virtuális környezet. Az adott batch fájl kétféleképpen is futtatható. Megkeressük a *Virtual\_ENV* mappában és jobb klikk után kiválasztjuk a „Futtatás adminisztrátorkánt” lehetőséget, majd ezt engedélyezzük a felugró ablakban. A második lehetőség, hogy indítunk egy terminált (command prompt) és a mappába navigálás után a *./setup\_virtual\_environment.bat* parancs kiadásával elindítjuk. Miután terminált a folyamat találunk négy virtuális gépet a rendszerünkön. Egy Mosquitto Brókert, ez fog a kommunikációért felelni, egy Control Node-ot, ami a folyamatot fogja irányítani és két felismerőt, Node1 és Node2, akik a felismerési folyamatért lesznek felelősek. Ezt tudjuk ellenőrizni a *multipass list* parancs parancssorból történő kiadásával.

A screenshot of a computer

Description automatically generated

1. ábra A multipass list parancs eredménye a virtuális környezet telepítése után

Az ábrán látszik, hogy létrehoztunk négy virtuális gépet. (1. ábra) A multipass list parancs eredménye a virtuális környezet telepítése után Az első oszlop tartalmazza a nevüket, a második, hogy aktívak és működnek, a harmadik az IP címüket és a negyedik, hogy milyen szoftver verzión futnak.

### Blueprint telepítése

A Blueprint telepítéséhez szükségünk van egy fájlra, ami függőleges vonalakkal (pipe szeparátor; |) elválasztva tárolja az összes Node felhasználónevét, IP címét és ha van jelszavát. Ennek a fájlnak a *Blueprint/nodelist.*txt útvonalon kell elhelyezkednie. Ennek az a funkciója, hogy a Blueprint-et telepítő szkript innen automatikusan beolvashassa az SSH-hoz szükséges adatokat, majd később, az alkalmazás futtatásakor újra felhasználja és ezeket ne a felhasználónak kelljen minden alkalommal manuálisan megadnia. Amennyiben virtuális környezetet használunk, ezt a szöveges fájlt az 1. ábra alapján kell kitölteni.

A screenshot of a computer screen

Description automatically generated

2. ábra nodelist.txt tartalma virtuális környezet esetén

Az első sorba kerülnek a Mosquitto Bróker adatai, a másodikba a Control Node adatai és ez után következik annyi felismerő Node adata, ahány a rendelkezésünkre áll. Ez a 2. ábrán kettő.

Miután a *nodelist.txt* szöveges fájlt kitöltöttük, a *Blueprint* mappában megtalálható a *setup.bat* batch fájl, ami felelős a Blueprint telepítéséért a megadott virtuális vagy fizikai eszközeinken. Amikor ezt futtatjuk, feltelepíti a megfelelő modulokat és Python könyvtárakat a megfelelő számítógépekre, majd terminál. Ennek a kódnak a futási ideje hosszúra nyúlhat, mivel több különböző nagyméretű telepítést hajt végre minden egyes felismerőn annak érdekében, hogy az alkalmazás akadály nélkül fusson.

### Zenefelismerő használata

A zenefelismerő alkalmazást a Blueprint telepítése után lehet használni. Ezt olyan módon tehetjük meg, hogy futtatjuk az *Applications/full\_lampion.bat* batch szkriptet. Ekkor elindul a felhasználói felület, amely különböző kérdések alapján összegyűjti, hogy milyen hanganyagot, hogyan szeretnénk felismerni. Az első kérdés, hogy hány hangszerből áll a felismerni kívánt dallam. Ez látható a 3. ábrán.

A screenshot of a computer screen

Description automatically generated

3. ábra Hangszerek számának bekérése a felhasználótól

A hangszerek száma nem lehet több a felhasználó számára elérhető virtuális gépek számánál. Ennek az az oka, hogy egy felismerő egy szólamot tud egyszerre feldolgozni, így lényegében minden felismerőhöz hozzárendelünk egy dallamot. Ezután a felhasználónak meg kell adnia, az adott hangszerek paramétereit. Ezek a következők: a hangszer neve, a hangszer által játszott dallam skálája és a hangfájl útvonala a személyi számítógépen. A hangszer neve bármi lehet, viszont a skálája ennél specifikusabb. Van két alapértelmezett opció, viszont lehetőség van egyedi skála megadására is. Ha a skála kromatikus („chromatic”) akkor az [ 'A', 'A#', 'B','C', 'C#', 'D', 'D#', 'E', 'F', 'F#', 'G', 'G#'] hangokat, ha pentatonikus („pentatonic”) akkor a [ 'A', 'C', 'D', 'E', 'G'] hangokat és ha egyedi akkor a felhasználó által megadott hangokat fogja felismerni az alkalmazás. Miután egy hangszer minden paraméterét megadta a felhasználó, az alkalmazás visszakérdez meg egyszer, hogy erősítse meg, vagy vesse el a konfigurációt, mivel ezt nem lesz lehetőség a későbbiekben módosítani. Az alkalmazás addig ismétel egy kérdést, amíg arra helyes paramétert vagy olyan fájlt nem kap, ami létezik. Egy kromatikus skálájú cselló paraméterezése látható a 4. ábrán.

A screenshot of a computer program

Description automatically generated

4. ábra Hangszer paramétereinek beolvasása a felhasználótól

Miután a felhasználó minden hangszert megadott és elfogadott, a hanganyagok kiküldésre kerülnek a felismerőkhöz, amelyek feldolgozzák majd továbbítják a kész hanganyagot a Control Node-nak, aki összeállítja a végső kottát. A kész musicxml formátumúkotta megtekinthető a */home/vmuser/Szakdolgozat\_v2/Application/musescore* mappában.

A következőkben egy futási eredmény látható különböző paraméterekkel. Minden esetben megegyezett a népdal és a hangszer viszont változott a skála paraméter. Az 5. ábrán látszik, hogy amikor kromatikus paraméterrel ismerjük fel a dallamot, akkor a skála összes hangja elérhető a felismerő számára.

A sheet music with musical notes

Description automatically generated

5. ábra "chromatic" paraméterrel futtatott felismerő

A 6. ábrán megfigyelhető, hogy amennyiben pentatonikus skálát adunk meg bemeneti paraméterként, a felismert hanganyagban kizárólag a skála hangjai jelennek meg. Nevezetesen az A, C, D, E és a G.



6. ábra "pentatonic" paraméterrel futtatott felismerő

A „custom” kulcsszó használatával ez tovább finomítható. Ennek abban az esetben van értelme, ha a felismerés előtt pontosan tudjuk, hogy milyen hangokból áll a dallam. Ekkor tudjuk biztosítani, hogy a kottában valóban csak azok a hangok jelenjenek meg amelyeket elvárunk.

## Futás közben adódó tipikus hibák megoldása

# Fejlesztői dokumentáció

## A probléma specifikálása

A mesterséges intelligencia alapú felvételről történő zenefelismerés egy kifejezetten körülményes, lassú és legfőképpen nehéz feladat. Telepíteni kell hozzá egy megfelelő környezetet, le kell tölteni a hozzá szükséges modulokat, választani kell egy programozási nyelvet, abban implementálni az alkalmazást, hiba és kompatibilitás ellenőrizni és futtatni, ami a mai, forgalomban lévő személyi számítógépeken legjobb esetben is körülményes, mivel egy Virtuális Géphez, amin TensorFlow fut, a minimum ajánlott RAM 8Gb. (EITCA, 2024) A Trend Force piackutató cég szerint a szakdolgozat írásakor (2024 Q3 - Q4) az átlagos, személyi számítógépekbe helyeztt RAM 11.8 Gb, ami nem elegendő két párhuzamosan futó felismerő megfelelő működéséhez. (trendforce.com, 2024) A szakdolgozatom ezekre a problémákra ajánl megoldást, egy olyan Blueprint (3.6) formájában, ami bizonyos előfeltételek teljesítése után képes egy kattintásra felépíteni egy olyan MQTT hálózatot, ahol a személyi számítógép helyett virtuális, vagy fizikai gépeken, párhuzamosan történik a felismerés. Ezek a gépek rendelkezhetnek olyan erőforrásokkal melyek nagyban meghaladják a személyi számítógép teljesítményét, ezáltal felgyorsítva a felismerési folyamatot. Jelen esetben egy olyan Blueprint-re van szükségünk, amely a számára elérhető fizikai vagy virtuális gépeken létrehoz egy Brókert, egy Control node-ot és legalább egy felismerőt. (1. ábra: Az alkalmazás MQTT hálózatának felépítése) Ez mindegyikre telepíti az MQTT-hez és az alkalmazás futtatásához szükséges modulokat, majd konfigurálja azokat és terminál. Ez után már futtatható lesz az alkalmazás.

## MQTT

Az MQTT egy üzenet küldő protokoll, ami a dolgok internetéhez (IOT) lett kifejlesztve. A fizikai architektúrája kliensekből áll, amelyek feliratkoznak vagy publikálnak egy témára és egy brókerből, aki megkapja az üzeneteket, témákba rendezi őket és továbbítja azoknak a klienseknek, akik feliratkoztak az adott témára. Az üzenetek mérete minimális, aminek következtében távoli gépeket tud összekötni minimális erőforrás és sávszélesség felhasználásával. Az alkalmazásunkban egy bróker köt össze több klienst. Ezek a kliensek a felhasználó, aki elindítja a felismerési folyamatot, a felismerő node-ok, akik felismerik a hanganyagokat és egy Control node, aki összegyűjti a beérkező szólamokat, kottába rendezi őket majd eltárolja a végeredményt.

A diagram of a computer

Description automatically generated

7. ábra: Az alkalmazás MQTT hálózatának felépítése

Mivel téma alapú a publikációs séma és egy témára bárhány kliens publikálhat, így a feldolgozott szólamokat eljuttatni a control node-nak egy kis sávszélességet igénylő, alacsony rendszerterhelésű feladat. Ezáltal a felismerő node-oknak, amelyek gyakran limitált erőforrású virtuális gépeken futnak, több erőforrása marad az alkalmazás futtatására, mintha például egy tradicionálisabb HTTP protokollon történne a kommunikáció. (emqx, 2024) A hálózathoz csatlakozó eszközök száma nincs meghatározva, ami előnyös egy olyan alkalmazás esetében, ami a feladatokat több felismerő között osztja szét, mivel a kommunikációs protokoll nem szab fizikai korlátot a node-ok lehetséges számának. A beépített QoS biztosíthatja az üzenetek megbízható megérkezését. Ennek három szintje van. 0, amikor egy üzenet maximum egyszer lesz elküldve, 1, amikor egy üzenet legalább egyszer lesz elküldve és 2, amikor az üzenet biztosan egyszer lesz elküldve.

## Crepe

A CREPE egy monofónikus hangmagasság meghatározó, amely mély konvolúciós neurális hálózaton alapul és közvetlenül idő-doménű hanghullám bemeneten dolgozik. A konvolúciós hálót többféle zenei hangon tanították, jól teljesít zajos környezetben, nyílt forráskódú és képes valós időben hangot felismerni. Ezen tulajdonságok ideálissá teszik egy olyan alkalmazáshoz, ahol ugyanannak a Python kódnak kell felismernie különböző hangszerekből származó, valószínűsíthetően zajos, koncertfelvételeket. (Kim, Salamon, Peter, & Pablo Bello, 2018)

## Python használata a Blueprint és az alkalmazás megírásához

A Python egy magas szintű programozási nyelv, ami egyre populárisabb a fejlesztők körében. Ezen kívül egy gazdag és jól karban tartott könyvtárral rendelkezik, amelyben szinte bármilyen feladathoz található egy annak elvégzésére való modul. Ennek következtében egy Blueprint fejlesztése közben nincs szükség arra, hogy más nyelven fusson a felismerő, a bróker, a control panel és a kommunikációért felelős script, ezek mind használhatják ugyanazt a nyelvet. Ez megkönnyíti a telepítést mivel a különböző virtuális gépeken egyedül Pythonra van szükség, ami helytakarékos. A fejlesztés szempontjából is előnyös a keresztplatform támogatása miatt. A Windows operációs rendszeren megírt Python kód futtatható Linuxon és fordítva.

## Virtuális környezet (*Virtual\_ENV/…*)

Az alkalmazás fejlesztése és tesztelése virtuális környezetben történt. Ennek több oka is volt. Az első és legfontosabb, hogy a valóságban távoli gépeken, szervereken futó alkalmazások leggyakrabban virtuális gépeken futnak, annak érdekében, hogy a rendszer erőforrásai megfelelően legyenek megosztva, illetve, hogy amennyiben egy alkalmazás futása közben probléma lép fel, az ne befolyásolja a többi folyamatot. A Facebook példának okáért Amazon Web Services virtuális gépein futtat alkalmazásokat, ilyen többek között a website futtatása és a felhasználói adatok tárolása. (Yadavpuneet, 2023) Mivel az alkalmazásunk alapból virtuális gépekre készül, így valódi szerverekre történő telepítéskor nem igényel további konfigurációt. A második, hogy a teljes MQTT hálózat és a rajta futó alkalmazás futtatható és tesztelhető legyen egy személyi számítógépen. Ez mind a két folyamatot megkönnyíti, mivel nem kell fizikai hardvert vásárolni, a létrehozott virtuális gépeket gyorsan lehet módosítani, optimalizálni mind RAM, mind háttértár szempontjából. További előnye a módszernek, hogy megtartja a fizikai szeparációt a gépek között, így a környezet hasonló marad egy valódi fizikai gépeken létrehozott MQTT hálózathoz.

### Multipass

A multipass egy platform független eszköz, amivel felhő stílusú virtuális gépeket lehet létrehozni. (Canonical Multipass, 2024) Azon kívül, hogy támogatja a személyes konfigurációt cloud-init interfészen keresztül, egy nagy kép könyvtárral érkezik, amiből szabadon lehet különböző célokra virtuális gépet választani. Erre egy kiváló példa a Mosquitto appliance, ami egy Mosquitto brókerként működő, előre elkészített virtuális gép. Olyan VM-eket hoz létre, amelyek nem befolyásolják a fizikai gépet, ennek következtében kiválóan lehet velük modellezni egy valós alkalmazás fizikai szerveren történő működését.

### SSH kulcsok generálása

A virtuális környezet létrehozásának első lépése egy SSH kulcs generálása. Ennek az a funkciója, hogy jelszó nélkül be lehessen lépni távolról a Node-okba. Ez azért fontos, mivel a Blueprint telepítése közben a fizikai gépnek be kell tudnia lépni a virtuális gépekbe, annak érdekében, hogy különböző folyamatokat elindíthasson, mint például a megfelelő fájlok letöltése GitHub-ról, vagy az MQTT működéséhez szükséges hálózati engedélyek beállítása. Az SSH kulcs generálása a következő kódrészlettel történik. A -C opcióval hozzáadunk egy vmuser kommentet. Ez lesz a felhasználónév, amit az SSH-hoz használni fogunk, a -f opcióval pedig megadjuk, hogy a kulcs milyen nevű fájlban legyen tárolva.

ssh-keygen -C vmuser -f multipass-ssh-key

### Cloud-Init.yaml fájl létrehozása

A Cloud-init egy standardizált megközelítés a cross-platform virtuális gépek konfigurációjának megadására. A Multipass, képes .yaml fájlból beolvasott konfiguráció alapján létrehozni egy virtuális gépet, így amennyiben átadjuk az általunk generált SSH kulcsot, azzal lehetővé tesszük, hogy jelszó azonosítás nélkül beléphessünk a kívánt virtuális gépbe.

users:

  - default

  - name: vmuser

    sudo: ALL=(ALL) NOPASSWD:ALL

    ssh\_authorized\_keys:

    - <ssh key>

A .yaml fájl az users kulcsszóval definiálja azokat a felhasználókat, amelyeket létre kell hozni a gépen. A -default azt jelzi, hogy alapértelmezett felhasználót kell beállítani ami a multipass-al létrehozott virtuális gép esetében Ubuntu. A -name adja meg az új felhasználó nevét. Ennek a névnek meg kell egyeznie az ssh kulcs generálásakor megadott névvel, vagyis amennyiben a „vmuser” nevet adtuk meg kommentben a -C opcióval az ssh-keygen-nek, akkor most is a „vmuser”-t kell használni. A sudo: ALL=(ALL) NOPASSWD:ALL jelszó nélküli sudo jogot ad a felhasználó számára, ami azt jelenti, hogy bármelyik felhasználó nevében futtathat parancsokat a rendszeren jelszó megadása nélkül. Az ssh\_authorized\_keys tartalmazza a publikus kulcsot, amely megadásával a felhasználó ssh hozzáférést nyerhet a virtuális géphez. Amennyiben az ssh kulcsot a -f multipass-ssh-key opcióval definiáltuk, akkor a multipass-ssh-key.pub fájl tartalmát kell beilleszteni ide.

### A virtuális gépek létrehozása és konfigurálása (*setup\_virtual\_env.bat*)

A virtuális gépek létrehozása és konfigurálása a *Virtual\_ENV/setup\_virtual\_env.bat* batch fájl futtatásával történik.

#### A broker node létrehozása (mosquitto\_broker\_setup.sh)

A multipass-ban egy előre elérhető appliance a mosquitto ubuntu alapú virtuális gép, amely indítás után képes üzeneteket küldeni és fogadni az MQTT protokollon keresztül. Először ellenőrizzük, hogy már létezik-e a Mosquitto virtuális gép, és amennyiben nem, létrehozzuk. A -n opcióval lehet nevet adni a VM-nek. Ellenkező esetben elindítjuk az előre létrehozott példányt.

multipass list | findstr /C:"MosquittoBroker" >nul

if %errorlevel% neq 0 (

    multipass launch appliance:mosquitto -n MosquittoBroker

    …

) else (

    multipass start MosquittoBroker

)

A bróker bár fut és pingelhető, a feladatát még nem képes ellátni, mivel nincs alapértelmezetten definiálva a hallgató port, nem biztos, hogy az összes IP címre hallgat.

#!/bin/bash

sudo cp /var/snap/mosquitto/common/mosquitto\_example.conf /var/snap/mosquitto/common/mosquitto.conf

echo "listener 1883 0.0.0.0" | sudo tee -a /var/snap/mosquitto/common/mosquitto.conf > /dev/null

echo "allow\_anonymous true" | sudo tee -a /var/snap/mosquitto/common/mosquitto.conf > /dev/null

sudo systemctl restart snap.mosquitto.mosquitto.service

echo "ubuntu:pwd" | sudo chpasswd

Ezeknek a problémáknak a kijavítására bemásoljuk a mosquitto\_example.conf fájlt a mosquitto.conf helyére, amivel alapértelmezett konfigurációt hozunk létre a Mosquitto számára. Ezután hozzáfűzzük a *listener 1883 0.0.0.0* sort a mosquitto.conf-hoz. Ez megadja, hogy a bróker az összes IP címre hallgasson a 1883-as porton. Engedélyezzük az anonim hozzáférést az *allow\_anonymus true* konfigurációs fájlba másolásával, hogy a brókerhez való csatlakozás során ne kelljen a node-oknak azonosítani magukat. Ez olyan megfontolásból történt, hogy semmilyen nagy kockázatú adat nem megy keresztül a hálózaton, így a biztonsági kockázat elhanyagolható. Ahhoz, hogy a módosítások érvénybe lépjenek, újra kell indítani a Mosquitto szolgáltatást, ami a *sudo systemctl restart snap.mosquitto.mosquitto.service* parancs futtatásával lehetséges. A Mosquitto kommunikációhoz, bár nem kapcsolódik, viszont hasznos itt beállítani, az alapértelmezett felhasználó jelszavát. Ez arra szolgál, hogy a későbbiekben, amennyiben szükséges be lehessen lépni SSH kapcsolattal a virtuális gépre. Az előbb említett konfigurációs lépéseket a mosquitto\_broker\_setup.sh shell szkriptben találjuk, amelyet a bróker VM indítása után rámásolunk, engedélyt adunk a futására, majd futtatjuk.

multipass transfer mosquitto\_broker\_setup.sh MosquittoBroker:/home/ubuntu/mosquitto\_broker\_setup.sh

    multipass exec MosquittoBroker -- chmod +x /home/ubuntu/mosquitto\_broker\_setup.sh

    multipass exec MosquittoBroker -- /home/ubuntu/mosquitto\_broker\_setup.sh

#### A Control Node és felismerő node-ok létrehozása

A virtuális környezet létrehozásakor a különböző node-oknak nincs semmiféle alapvető feladata azon kívül, hogy az előfeltételeknek megfelelő konfigurációval induljanak el. Ez ebben az esetben azt jelenti, hogy lehessen rá telepíteni a megfelelő modulokat és lehessen vele SSH kapcsolatot létesíteni. Bizonyos Python könyvtárak használata esetében ez különösen fontos, mivel például a Tensorflow-hoz, ami kritikus része a zenefelismerő alkalmazás működésének, az ajánlott minimális memória 8 GB. (EITCA, 2024) A személyi számítógépemen erre nincsen lehetőség, mivel ez már két hangszer esetén is több memóriát igényelne, mint amennyi rendelkezésre áll. Az előbb említett limitáció következtében kénytelen voltam kikísérletezni „trial and error” alapján, a minimálisan szükséges RAM-ot, ami bár nem elég az optimális működéshez, de megfelelő tesztelésre. Ez a szám a lapozási mechanizmus engedélyezését követően két gigabájtra jött ki. Az előbbiekben felsorolt kritériumokat figyelembe véve a node-ot a következőképpen kell elindítani:

multipass launch jammy -n %node1% --cloud-init cloud-init.yaml --disk 10G --memory 2G

Ahol a -n kapcsoló kell megadni a virtuális gép nevét, a –cloud-init után a 3.5.3 szekcióban említett konfigurációs fájlt, a –disk után a háttértár méretét és a –memory után a memória méretét.

## Blueprint (*Blueprint/…*)

A „Blueprint” azaz tervrajz általában egy olyan vázlatot jelent az informatikában, amely egy rendszer, szoftver vagy alkalmazás struktúráját és működését írja le. Ez a gyakorlatban annyit jelent, hogy egy ismeretlen, előre nem definiált környezetben, automatikusan felállítja egy adott alkalmazás működéséhez szükséges kommunikációs hálót és telepíti a rendszer főbb komponensein a megfelelő modulokat és eszközöket, majd konfigurálja azokat.

Léteznek ugyan nyílt forráskódú, mindenki számára elérhető blueprintek, mint például az Akarino projekt, viszont ezek a szoftverek nagyon ritkán vannak frissítve, aminek következtében, a használatuk legjobb esetben is körülményes. (Thorking, 2024) Bizonyos bennük használt modulok mára már elérhetetlenek, vagy teljesen máshogy működnek, mint ahogyan eredetileg. Ezt mutatja be a második ábra, ahol különböző Akraino Blueprintek kerültek összehasonlításra adott szempontok alapján. Észrevehető, hogy bizonyos Blueprintekből a legutolsó kiadás is legalább kétéves a szakdolgozat írásának pillanatában, ezen felül az adott verzió dokumentációja csak egy referencia egy régebbire ami akár 6 – 8 éves is lehet és javarészt elavult. (2. ábra: Már létező Akrainó Blueprintek összegzése különböző paraméterek szerint) Ennek következtében amennyiben egy kommunikációs hálót, vagy egyedi igényű futási környezetet szeretnénk Blueprint segítségével implementálni, jobban járunk, mind idő, mind energiabefektetésileg, ha megírjuk a sajátunkat.



8. ábra: Már létező Akrainó Blueprintek összegzése különböző paraméterek szerint

Bizonyos technikák azonban elleshetőek és univerzálisan használhatóak az előbb említett konstrukciókból. Ilyen például a nodelist.txt ami egy pipe („|”) szimbólummal elválasztva tartja számon az összes komponens felhasználónevét, IP címét és amennyiben van, jelszavát. Ennek a fájlnak a létezése nagyban elősegíti a skálázhatóságot mivel azzal szemben, hogy minden node-ra egyesével telepítenénk a kívánt modulokat, csak egy ciklus segítségével végig iterálunk a fájlon és végrehajtjuk a kívánt műveleteket minden elemen.

### Control Node konfigurálása (*control\_node\_setup.sh*)

A Control Node-nak a feladata, hogy fogadja a felismerők által feldolgozott adatokat, majd azokat feldolgozza és tárolja. Mivel az üzeneteket MQTT-n keresztül kapja, így szükség van olyan kliensprogramokra, illetve Python könyvtárakra, amelyek képessé teszik ezek küldésére és fogadására. Ilyen a mosquitto-clients és a paho-mqtt. A feldolgozni kapott adatokat musicxml formátumban kapja, és a Python programozási nyelv segítségével dolgozza fel, így szükség van a music21 könyvtárra is, ami egy Python alapú eszköztár számítógépes zeneelmélethez. (Cuthbert, 2023)

Annak érdekében, hogy a telepítési folyamat skálázható legyen akár távoli, vagy fizikai szerverekre a multipass beépített shell parancsa helyett, ami csak lokálisan működik, SSH-val lépünk be a node-ba. A nodelist szöveges fájlból beolvassuk a Control node IP címét és a hozzá tartozó felhasználónevet, majd a Windows-ba beépített SSH parancs segítségével belépünk távolról a node-ba. (Dows, 2024) Ezután, ha már létezik töröljük, majd GitHub-ról letöltjük a Blueprint legújabb változatát, ami tartalmazza mind a telepítőket, mind a futtatható alkalmazásokat. Ezután belépünk a Blueprint mappába, engedélyt adunk a control\_node\_setup.sh-nak a futásra és futtatjuk. Ez telepíti az előző bekezdésben említett szükséges kliensprogramokat és Python könyvtárakat, majd létrehozza a messages mappát és terminál.

ssh %%a@%%b -i ../Virtual\_ENV/multipass-ssh-key "if [ -d 'Szakdolgozat\_v2' ]; then rm -rf Szakdolgozat\_v2; fi && git clone %REPO\_URL% && cd Szakdolgozat\_v2 && chmod +x Blueprint/control\_node\_setup.sh && ./Blueprint/control\_node\_setup.sh"

### Node-ok konfigurálása (*node\_setup.sh*)

A felismerő node-ok konfigurálása hasonló a control node-hoz abból a szempontból, hogy a nodelist szöveges fájlból kiolvassuk a node IP címét és felhasználónevét majd ssh kapcsolattal belépünk rá, letöltjük GitHub-ról a blueprintet majd végrehajtjuk a node\_setup.sh-t. Azonban node-okból bármennyi lehet, aminek következtében az előbb említett folyamatot beágyazzuk egy ciklusba, ami végig iterál a hátralévő sorokon és mindegyikre végrehajtja a konfigurációs folyamatot. Ezen kívül a node-okon történik a zenék felismerése, python programozási nyelv segítségével, aminek következtében a control node-hoz szükséges a tensorflow a crepe működéséhez és a crepe a zenefelismeréshez. Továbbá, annak érdekében, hogy a node ne fogyjon ki a memóriából a tensorflow telepítése közben, a 3.5.4.2 szekcióban említettek alapján be kell állítani a lapozási mechanizmust.

#!/bin/bash

#lapozási mechanizmus

sudo fallocate -l 1G /swapfile

sudo chmod 600 /swapfile

sudo mkswap /swapfile

sudo swapon /swapfile

sudo apt install mosquitto-clients -y

sudo apt install python3-pip -y

sudo pip install --upgrade tensorflow

pip install crepe

pip install music21

pip install paho-mqtt

## Zenefelismerő alkalmazás (*Application*)

A zenefelismerő alkalmazás szétosztja a hangszereket és azok szólamát tartalmazó wav fájlokat a felismerő node-ok között, amelyek feldolgozzák és továbbítják az így keletkezett musicxml dokumentumot a Control node felé, aki ezeket egy végleges kottává olvasztja össze és tárolja. Ez után az alkalmazás terminál. Ezt a viselkedést olyan módon lehet elérni, hogy futtatjuk a *full\_lampion.bat* batch szkriptet. Ekkor automatikusan elindul a *lampion.py* amely bekéri a felhasználótól a hangszerek adatait, majd tovább küldi őket feldolgozásra. A következő ábra (3. ábra: Az alkalmazás híváslistája és a hívott szkriptek híváslistája) szemlélteti, hogy milyen folyamatok játszódnak le a *full\_lampion.bat* batch szkript futtatása közben. Először beolvasásra kerülnek a belépési adatok a *nodelist.txt*-ből, majd mind a Control Node-on, mind a felismerőkön végrehajtásra kerülnek az alkalmazás futásához szükséges szkriptek és utasítások. A mappák dinamikus létrehozásának és törlésének az a szerepe, hogy egy Node többszörösen is felhasználható legyen különböző bemenetekkel. Az ábrán említett Python szkriptek tartalma és feladata a következő fejezetekben kerül kifejtésre.

A black background with white text

Description automatically generated

9. ábra: Az alkalmazás híváslistája és a hívott szkriptek híváslistája

### Adatok bekérése a felhasználótól (*lampion.py*)

Az adatok bekérése a felhasználótól a *lampion.py* python szkript futtatásával történik. A *lampion.py* fájl egy Python szkript, amely különböző eszközök (instruments) és csomópontok (nodes) közötti kommunikációt valósít meg MQTT protokoll segítségével. A szkript két fő funkciót tartalmaz: az egyik a vezérlő csomópont (control node) inicializálása, a másik pedig az eszközök csomópontjainak indítása és a szükséges fájlok küldése. Ezeket a fájlokat a felhasználótól kéri be ellenőrzött módon. Először a hangszerek számát, majd az adott hangszerek nevét, skáláját majd azt, hogy hol található a felismerni kívánt wav fájl. Ezt követően a felhasználónak meg kell erősítenie a hangszer paramétereit, mivel ezt visszamenőlegesen nem lehet módosítani. Az így bekért adatokat egy *Instrument* típusú objektumban tárolja minden hangszerhez.

temp\_instrument = Instrument(instrument\_name, i, instrument\_scale, wav\_file\_location)

Az *initiate\_controll\_node* függvény felelős a vezérlő csomópont inicializálásáért. Ez a függvény egy alprocesszt indít, amely a *mqtt\_publish\_message.py* szkriptet futtatja, és egy üzenetet küld a vezérlő csomópontnak arról, hogy hány eszköz van csatlakoztatva. 3.7.2 A függvényt egy külön szálban (control\_thread) futtatják, hogy ne blokkolja a fő program futását.

def initiate\_controll\_node():

    subprocess.run(["python3", "mqtt\_publish\_message.py", broker\_ip, "ControlNode/initiate\_control\_node", str(len(instruments))])

A *start\_instrument\_node* függvény minden egyes eszközhöz és csomóponthoz külön-külön fut le. Először egy üzenetet küld az adott csomópontnak, hogy indítsa el az eszközt az előzőekben bemutatott módon. Ezután létrehoz egy ideiglenes fájlt, amely tartalmazza az eszköz nevét és skáláját, majd ezt a fájlt és az eszközhöz tartozó WAV fájlt is elküldi az adott csomópontnak. Végül egy üzenetet küld, hogy jelezze az eszköz inicializálásának végét.

A szkript több szálat használ a párhuzamos végrehajtáshoz. Minden eszközhöz és csomóponthoz létrehoz egy új szálat, amely a *start\_instrument\_node* függvényt futtatja. Ezeket a szálakat egy listában (*threads*) tárolja, és mindegyiket elindítja. A program végén minden szálat megvár (*join*), hogy biztosítsa, hogy minden eszköz inicializálása befejeződjön, mielőtt a vezérlő csomópont inicializálása is befejeződik.

for thread in threads:

    thread.join()

control\_thread.join()

### Üzenet és fájl küldése MQTT-n Python szkripttel (*mqtt\_publish\_message.py és mqtt\_publish\_file.py*)

Annak érdekében, hogy az alkalmazás minél könnyebben futtatható legyen különböző operációs rendszereken az MQTT-n történő üzenetek és fájlok küldése a felhasználó személyi számítógépén Python szkripttel lett megvalósítva. Az MQTT kliens kapcsolódik a szerverhez, majd közzéteszi az üzenetet és lecsatlakozik.

client.connect(broker, port)

# Start the network loop

client.loop\_start()

# Publish the message

result = client.publish(topic, message)

# Wait for the message to be published

result.wait\_for\_publish()

# Stop the network loop and disconnect

client.loop\_stop()

client.disconnect()

Fájl küldése esetén ez a folyamat annyiban módosul, hogy a fájl bináris módon kerül megnyitásra (rb) majd publikálásra.

with open(file\_path, 'rb') as file:  # Open the file in binary mode

        file\_content = file.read()

        client.publish(topic, file\_content)

### Control Node működése az alkalmazás futása közben (*control\_node*\_*action.sh*)

A Control Node az alkalmazás indulásakor nem tudja, hogy hány darab feldolgozott musicxml fájlt vár, amelyek az egyesítésre váró szólamokat tartalmazzák. Ezt a számot, amit a továbbiakban hívjunk „n”-nek ( n∈ ℕ+) akkor fogja megkapni, amikor a felhasználó már megadta és megerősítette a felismerni kívánt szólamokat. Ez után „n” darab musicxml fájlt vár, amiket ideiglenesen tárol, majd, amikor megérkezett az összes egyesít. Utolsó lépésben a kottát elmenti és törli a temporális fájlokat, majd terminál.

#### Hallgatódzó mód (listen\_control\_node.py)

A hallgatódzási folyamat a *listen\_control\_node.py* Python szkript futtatásával valósul meg. Először ellenőrizzük, hogy a könyvtár ahová menteni szeretnénk az ideiglenes fájljainkat létezik-e. Amennyiben nem, létrehozzuk. Ezután létrehozzuk a *callback* függvényeket azokra az esetekre, amikor csatlakozunk, vagy üzenetet kapunk. A csatlakozás *callback* függvénye kiírja, hogy milyen visszatérési értékkel sikerült csatlakozni a szerverhez, majd feliratkozik a *ControlNode/+* témára. A „+” jel, jelen esetben egy *wildcard* ami azt jelenti, hogy a *ControlNode* téma minden altémájára érkező üzenetet olvasson be.

def on\_connect(client, userdata, flags, rc):

    print(f"Connected with result code {rc}")

    client.subscribe(TOPIC)

A beérkezett üzenetek *callback* függvénye kétféle módon működik, amelyek között a „first” logikai változó igazra, vagy hamisra változtatásával tudunk kapcsolni. Az első üzenet előtt ez a logikai változó igaz, ami azt jelenti, hogy a *CotrolNode/* *+* témáról egy számot szeretnénk beolvasni. Ez a szám lesz az „n”, amely meghatározza, hogy hány darab feldolgozott szólamot kell a továbbiakban várni. Ez után a „first” változó értékét hamisra állítjuk. Ez után „n” darab fájlt olvasunk ki a *ControlNode/+* témából és a tartalmukat a messages mappába mentjük. Annak érdekében, hogy minden szólamnak a neve egyedi legyen, a beérkezett musicxml fájloknak a téma nevét adjuk, amelyen érkeztek. Ennek az egyediségét a felismerő node-ok biztosítják azáltal, hogy a saját felhasználónevüket használják publikálásnál altémának. (3.7.2.5) A kívánt számú beérkezett fájl után a hallgatódzó módért felelős python szkript lecsatlakozik az MQTT hálózatról és terminál.

#### A beérkezett feldolgozott anyagok egyesítése, tárolása és az ideiglenes fájlok törlése (unification.py)

A beérkezett musicxml fájlokat a *listen\_control\_node.py* a messages mappába írta, olyan módon, hogy a nevük egy egyedi azonosító, amely magába foglalja, hogy melyik node-ról érkeztek. Ezeket a fájlokat az *unification.py* segítségével beolvassuk, a music21 könyvtár segítségével ellenőrizzük, hogy csak egy szólamot tartalmaznak és amennyiben igen, akkor hozzáadja a „combined\_score” *stream.Score()* típusú változóhoz. Amikor a messages mappa összes musicxml fájlján végigért a szkript, a „combined\_score” változóban tárolt teljes kotta tartalmát a kimeneti a musescore mappába és terminál. Ezután töröljük a messages mappa tartalmát, hogy felkészüljünk a következő kombinálási feladatra.

### A felismerő node működése az alkalmazás futása közben (*node\_action.sh*)

A felismerő node feladata, hogy a beérkezett WAV formátumú hanganyagot a megadott paraméterek alapján, mint a hangszer és annak skálája, a Crepe mély neuronháló segítségével felismerje, feldolgozza, musicxml formátummá konvertálja, majd továbbítsa a Control Node felé az MQTT hálózaton keresztül. Mivel nem hozunk létre új node-okat minden feladathoz, így a node-oknak univerzális felismerőként kell funkcionálniuk, amelyek megfelelő paraméterezés mellett képesek feldolgozni különböző hangszerekből és stílusokból származó hanganyagokat.

#### Hallgatódzó mód (listen\_node.py $(username))

A node univerzális mivolta következtében szüksége van bizonyos paraméterekre a felismerési folyamat elvárt végrehajtásához. Ezek a paraméterek a hangszer neve, skálája és a felismerni kívánt WAV hangfájl. Ezeket az adatokat MQTT-n keresztül a felhasználótól kapja meg. A *node\_action.sh* amit az alkalmazás indításánál futtatunk a felismerő node-on (3.7) először létrehoz egy ideiglenes mappát, azon belül az *out.txt* temporális kimeneti fájlt, majd elindítja a *listen\_node.py*-t a *$(hostname)* bemeneti paraméterrel. Ez azért fontos, mivel ezen a témán fog hallgatni a felismerő bejövő üzenetekre, és mivel a node neve egy egyedi azonosító, ezzel tudjuk biztosítani, hogy ne keveredjenek össze az üzenetek. A node elkezd hallgatni a „*system/$(username)/+*”témán és addig vár, amíg nem kap egy „start” üzenetet. Amikor ezt megkapta, az ezután következő üzeneteket fájlként tárolja az ideiglenesen létrehozott mappába (TMP) addig amíg nem érkezik egy „end”, ami után terminál. Annak érdekében, hogy a beérkezett fájlok helyes formátumban legyenek tárolva, az eredeti nevén létrehozott témában publikáljuk a tartalmát. Ezután a témát adjuk a fájl nevének a felismerő node-on. Például a „beres.wav” hangfájlt a „*system/Node1/beres.wav”* témán publikáljuk, amit aztán a node beolvas és a „*TMP/system\_Node1\_beres.wav*” néven ment.

topic = msg.topic

            payload = msg.payload

            file\_path = os.path.join(SAVE\_DIR, topic.replace('/', '\_'))

            print(f"{topic} {file\_path}")

            with open(file\_path, 'wb') as f:

                f.write(payload)

#### Beérkezett hanganyag felismerése a paraméterek alapján (pitch\_detection.py TMP/\*.wav)

A mentett WAV fájl felismerése a *pitch\_detecion.py TMP/\*.wav* szkript végrehajtásával történik. Mivel az ideiglenes mappát a felismerési folyamat végén töröljük és csak egy audió fájlt kapunk node-onként, így biztosak lehetünk benne, hogy a *TMP* mappa egy .wav kiterjesztésű elemet tartalmaz. A *scipy.io* könyvtárat használva beolvassuk a mintavételezési rátát és a hanganyagot a *sr* és *audio* változókba.

filename = sys.argv[1]

sr, audio = wavfile.read(f"{filename}")

Ez után meghívjuk a *crepe*.*predict(...)* függvényt, paraméterül adjuk neki a hanganyagot, a mintavételezési rátát és bekapcsoljuk a viterbi simítást. A függvény egy rendezett négyessel tér vissza, ami tartalmazza az időt, a jósolt frekvenciát, a konfidenciát és a nyers aktivációs mátrixot. A *numpy* könyvtár használatával egy kétdimenziós mátrixba rendezzük a rendezett négyes első három elemét, mivel az aktivációs mátrixra nem lesz szükség a feldolgozáshoz. Ez után az időt kicseréljük mintavételezési rátára, olyan módon, hogy a létrehozott *numpy.colum\_stack* első oszlopát egyenlővé tesszük két egymást követő idő különbségével.

time, frequency, confidence, activation = crepe.predict(audio, sr, viterbi=True)

data = numpy.column\_stack((time, frequency, confidence))

rate = data[1][0] - data[0][0]

for i in range(len(data)):

    data[i][0] = rate

Ez után a nyers felismerési adatot mentjük a *TMP/out.txt* fájlba tabulátorokkal elválasztva, 3 tizedesjegyre kerekítve, majd a program terminál.

numpy.savetxt('TMP/out.txt', data, fmt='%.3f', delimiter='\t')

#### Felismert hanganyag feldolgozása (process.py TMP/system\*.txt)

Miután a *pitch\_detection.py* terminál a *client\_action.sh* meghívja a *process.py* python szkriptet, aminek a feladata, hogy feldolgozza a nyers adatokat. Először létrehoz egy tömböt a lehetséges frekvenciákkal egy oktávon. Ez olyan módon történik, hogy alapul vesz egy kromatikus skálát, ahol az „A” hang felel meg a 440Hz-nek majd kiszámolja, hogy a bizonyos hangok hol helyezkednek el a 0Hz-től számított első 8 oktávban. Ez a lista lesz az, ami az elérhető hangokat tárolja. Ennél azért nincs szükség szélesebb frekvencia halmazra, mivel az emberi fül ennél magasabb frekvenciát már nem hall. (Pumphrey, 1950)

notes = [ 'A', 'A#', 'B','C', 'C#', 'D', 'D#', 'E', 'F', 'F#', 'G', 'G#']

octaves = np.arange(0, 9)

frequencies = []

for octave in octaves:

    for note in notes:

        frequency = 440 \* (2 \*\* (octave - 4 + (notes.index(note) / 12)))

        frequencies.append(frequency)

note\_freq\_array = np.array(frequencies)

note\_freq\_array\_2 = list(zip(9\*notes,note\_freq\_array))

note\_freq\_array = np.append(note\_freq\_array,0)

Ez így készített skála kromatikus, ez viszont nem minden esetben optimális. Ez annak a következménye, hogy vannak olyan darabok, ahol ennek csak egy részhalmazát használjuk. A 3.7.1-ben láttuk, hogy a felhasználótól bekérjük a hangszer skáláját. Ez az információ itt kerül felhasználásra. Amennyiben a felhasználó a „chromatic” kulcsszót használta, nem történik semmi. A „pentatonic” kulcsszó használatával a skála leszűkül az „A,C,D,E,G” ötfokú, azaz pentatónikus, részhalmazra. Amennyiben a felhasználó a „custom” kulcsszót használta és egyedi skálát adott meg, abban az esetben csak azok a hangok fognak az eredményben megjelenni, amelyeket megadott.

if content[1].upper() == "chromatic".upper():

    pass

elif content[1].upper() == "pentatonic".upper():

    t\_notes = [ 'A', 'C', 'D', 'E', 'G']

    note\_freq\_array\_2 = [item for item in note\_freq\_array\_2 if item[0][0:] in t\_notes]

    note\_freq\_array = np.array([freq for note, freq in note\_freq\_array\_2])

else:

    t\_notes = content[1].strip('[]').replace("'", "").split(',')

    note\_freq\_array\_2 = [item for item in note\_freq\_array\_2 if item[0][0:] in t\_notes]

    note\_freq\_array = np.array([freq for \_, freq in note\_freq\_array\_2])

Ez után beolvassuk a felismerés végén tárolt fájlt és elkezdjük a feldolgozást. Az adatokat egy *numpy* tömbbe tároljuk el. Amennyiben egy felismert frekvenciának a konfidenciája 0.5 alatt van akkor egyenlővé tesszük a hozzá legközelebbi értékkel. Ezután megkeressük az érvényes frekvenciák közül, hogy melyik van a legközelebb az adott méréshez egyenlővé tesszük. Például, ha a felismerő 439.83 Hz-t mért, akkor ezt lecseréljük a hozzá abszolút értékben legközelebb lévő frekvenciára, a 440 Hz-re.

def find\_closest\_frequency(number):

    absolute\_diff = np.abs(note\_freq\_array - number)

    closest\_index = np.argmin(absolute\_diff)

    closest\_frequency = note\_freq\_array[closest\_index]

    return closest\_frequency

Annak érdekében, hogy megkapjuk, mennyi ideig tartott egy hang az időben, összegezzük az egymás után következő azonos frekvenciájú mérések mintavételezési idejét, és átlagoljuk a konfidenciájukat.

count=0

sum = 0

confidence = 0

last\_value = -1

data\_2 = np.array([])

for i in range(0, len(data)):

    value = data[i,1]

    time = data[i,0]

    confidence\_tmp = data[i,2]

    if value == last\_value or last\_value == -1:

        sum+=time

        confidence += confidence\_tmp

        count += 1

        last\_value = value

    else:

        data\_2 = np.append(data\_2,[last\_value, sum, (confidence / count)],axis = 0)

        count=0

        sum = 0

        confidence = 0

        last\_value = -1

A jelenlegi adatban még előfordulhatnak felismerési hibák, olyan hangok amelyek valamilyen külső vagy belső tényező következtében tévesen lettek mérve. Ezeknek a soroknak két jellemzője van, irreálisan rövid ideig tartanak (három századmásodpercnél rövidebbek), vagy nagyon alacsony átlagos konfidenciaszinttel rendelkeznek. Annak érdekében, hogy ne rövidítsük meg a sorokat, mivel így elveszne a szólamok közötti összhang, az ilyen hibás adatok helyére szünetet illesztünk be. A feldolgozott adatot, ami a kezdeti mintavételezési idő, frekvencia, konfidencia helyett most már a hang hossza, hangmagasság, átlagos konfidencia sorokat tartalmazza elmentjük a *TMP/processed.txt* szöveges fájlba az utolsó oszlopot elhagyva, mivel arra a továbbiakban nem lesz szükségünk.

to\_delete = []

for i in range(1,len(data\_2)-1):

    if((data\_2[i, 1] < 0.1) and (data\_2[i-1, 0] == data\_2[i+1, 0])):

        …

    elif((data\_2[i, 1] < 0.05)): … #0.1 > s

    elif(data\_2[i, 2] < 0.70 and round(data\_2[i,0]) != 0):

         …

#### A feldolgozott hanganyag konvertálása musicxml formátummá (convert\_to\_musescore.py)

A feldolgozott fájlt, a *processed.txt*-t átadjuk a *convert\_to\_musescore.py* Python szkriptnek, ami a music21 könyvtár segítségével musicxml formátumúvá konvertálja a felismert adatokat. A hanghosszokat, a *processed.txt* első oszlopát, zeneileg értelmezhető hosszokra konvertálja. Ezek közül a legkisebb a másodperc egy-harmincketted része (0.03125s), ennek következtében töröltünk minden mérést, ami három századmásodpercnél rövidebb volt. 3.7.2.3

durations = [1/32, 1/16, 1/8, 1/4, 1/2, 1, 2, 4, 8, 16]  # Available durations

closest\_duration = min(durations, key=lambda x: abs(x - timestamp))

Ez után létrehozunk egy music21 *stream.Score()* objektumot amelybe beleírjuk a hangokat vagy amennyiben 0 volt a frekvencia a szüneteket és a hozzájuk tartozó időintervallumokat, majd az eredményt musicxml formátumban mentjük.

def generate\_musicxml(durations, frequencies, output\_file):

    # Create a music21 Stream object to store the musical elements

    score = stream.Score()

    # Create a Part object to hold the notes

    part = stream.Part()

    score.append(part)

    # Iterate over the durations and frequencies arrays

    for dur, freq in zip(durations, frequencies):

        if freq == 0:

            # Create a Rest object for the pause

            r = note.Rest()

            # Create a Duration object with the corresponding duration value

            d = duration.Duration(dur)

            # Set the quarter length of the Rest object

            r.quarterLength = d.quarterLength

            # Add the Rest object to the Part object

            part.append(r)

        else:

            # Create a Note object with the corresponding frequency

            n = note.Note()

            n.pitch.frequency = freq

            # Create a Duration object with the corresponding duration value

            d = duration.Duration(dur)

            # Set the quarter length of the Note object

            n.quarterLength = d.quarterLength

            # Add the Note object to the Part object

            part.append(n)

    # Convert the score to MusicXML format and save to file

    score.write('musicxml', fp=output\_file)

#### Az eredmény publikálása a Control Node felé (mosquitto\_pub ...)

A felismerő node miután felismerte a beérkező hangfájlt, feldolgozta és tárolta musicxml formátumban közzé teszi ezt a saját témáján az MQTT hálózaton annak érdekében, hogy a Control Node beolvashassa és feldolgozhassa. 3.7.1.1 Ezt a mosquitto\_pub parancssori eszköz segítségével hajtjuk végre. A -h kapcsoló meghatározza a bróker elérési címét, a -t a téma nevét és a -f a küldésre szánt fájl helyét.

mosquitto\_pub -h MosquittoBroker -t "ControlNode/$hostname" -f TMP/\*.musicxml

A „\*” helyettesítő karakter a *\*.musicxml* fájlnévben olyan megfontolásból lett használva, hogy a shell szkriptet ne kelljen node-specifikussá tenni. Minden node egy musicxml fájlt hoz létre, viszont ezeket nem biztos, hogy azonos néven mentik el. A „\*” használatával kiküszöbölhető minden lehetséges ebből adódó probléma.

#### Az ideiglenes fájlok törlése

Annak érdekében, hogy az alkalmazás következő futásakor ne lépjen fel konfliktus a régi és az új fájlok között, a *node\_action.sh* shell szkript végén az ideiglenes mappát töröljük.

rm -rf TMP

Ezzel véget ér a felismerési és publikálási folyamat és a *node\_action.sh* terminál.

## A felhasznált modulok és könyvtárak listája és verziószáma

## Tesztelési terv és a tesztelés eredményei

# Összefoglalás és további fejlesztési lehetőségek

# Irodalomjegyzék

Cuthbert, M. (2023). *music21.org*. Forrás: music21 Documentation: https://www.music21.org/music21docs/index.html

Dows, N. (2024. 05 30). *learn.microsoft.com*. Forrás: Tutorial: SSH in Windows Terminal: https://learn.microsoft.com/en-us/windows/terminal/tutorials/ssh

*EITCA*. (2024. 10 18). Forrás: What is the minimum amount of RAM recommended for allocating to the virtual machine running TensorFlow?: https://eitca.org/artificial-intelligence/eitc-ai-dltf-deep-learning-with-tensorflow/tensorflow/installing-tensorflow/examination-review-installing-tensorflow/what-is-the-minimum-amount-of-ram-recommended-for-allocating-to-the-virtual-machine-running-tens

*EITCA Academy*. (2023. 08 08). Forrás: What is the minimum amount of RAM recommended for allocating to the virtual machine running TensorFlow?: https://eitca.org/artificial-intelligence/eitc-ai-dltf-deep-learning-with-tensorflow/tensorflow/installing-tensorflow/examination-review-installing-tensorflow/what-is-the-minimum-amount-of-ram-recommended-for-allocating-to-the-virtual-machine-running-tens

*emqx*. (2024. 9 24). Forrás: MQTT vs HTTP: Ultimate IoT Protocol Comparison Guide: https://www.emqx.com/en/blog/mqtt-vs-http

Kim, J. W., Salamon, J., Peter, L., & Pablo Bello, J. (2018). *arxiv.org*. Forrás: CREPE: A Convolutional Representation for Pitch Estimation: https://arxiv.org/abs/1802.06182

Pumphrey, R. (1950). Upper Limit of Frequency for Human Hearing. *Nature*, 166.

Thorking. (2024. 10 18). *lf-akraino.attlassian.net*. Forrás: Akraino Wiki: https://lf-akraino.atlassian.net/wiki/spaces/AK/overview?homepageId=13664258

*trendforce.com*. (2024. 06 25). Forrás: CSPs to Expand into Edge AI, Driving Average NB DRAM Capacity Growth by at Least 7% in 2025, Says TrendForce: https://www.trendforce.com/presscenter/news/20240625-12200.html

Yadavpuneet. (2023. 09 11). *medium.com*. Forrás: Why Does Facebook Use Amazon Web Services (AWS) for its Infrastructure?: https://medium.com/@yadavpuneet2001/why-does-facebook-use-amazon-web-services-aws-for-its-infrastructure-dc9759d44134#:~:text=Scalability%20and%20Global%20Reach,to%20users%20across%20diverse%20regions.

# Melléklet