**ບົດໂຄງການຈົບຊັ້ນລະດັບປະລິນຍາຕີວິທະຍາສາດ**

**ສາຂາ ວິທະຍາສາດຄອມພິວເຕີ**

**ການຮຽນຮູ້ ແລະ ການຈື່ຈຳ ຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ ດ້ວຍ****ເຕັກນິກຂອງປັນຍາປະດິດຂັ້ນສູງ**

**Lao Text To Speech Recognition** **Using Deep Learning Technology**

**ນັກສຶກສາ:**

**ທ້າວ ຄຳປະເສີດ ໄຊຍະວົງ**

**ທ້າວ ໃຫຍ່ ແສງວິໄຊ**

**ອາຈານຜູ້ນຳພາ:**

**ອຈ.ປອ. ລັດສະໝີ ຈິດຕະວົງ**

**ອາຈານຜູ້ຊ່ວຍນຳພາ:**

**ອຈ.ປທ. ສົມມິດ ທຸມມາລີ**

**ຄະນະວິທະຍາສາດທຳມະຊາດ**

**ພາກວິຊາ ວິທະຍາສາດຄອມພິວເຕີ**

**ສົກສຶກສາ: 2022-2023**

**Final Project of Bachelor Degree of Natural Science in**

**Computer Science**

**ການຮຽນຮູ້ ແລະ ການຈື່ຈຳ ຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ ດ້ວຍເຕັກນິກຂອງປັນຍາປະດິດຂັ້ນສູງ**

**Lao Text To Speech Recognition Using** **Deep Learning Technology**

**Students:**

**Mr. Khampaserth XAIYAVONG**

**Mr. Gnay SAENGVIXAI**

**Advisor:**

**Lathsamy CHIDTAVONG, Ph.D**

**Co-Advisor:**

**Mr. Sommith THOUMMALY**

**Faculty of Natural Sciences**

**Department of Computer Science**

**Academic Year 2022-2023**

ຄຳນຳ

ໃນຍຸກທີ່ໂລກກຳລັງໝຸນໄປທາງໜ້າເອີ້ນໄດ້ວ່າແມ່ນຍຸກຂອງດິຈິຕອລ, ທຸກອົງກອນບໍ່ວ່າ ຈະເປັນພາກລັດ ຫຼື ເອກະຊົນນັ້ນລ້ວນແຕ່ໃຊ້ເຕັກໂນໂລຊີບໍ່ທາງໃດກໍທາງໜຶ່ງເຂົ້າມາຊ່ວຍໃນ ວຽກງານຕ່າງໆ ເພື່ອໃຫ້ຄວາມສະດວກສະບາຍ, ຄວາມວ່ອງໄວ ແລະ ຄວາມແມ່ນຢໍາໃນການ ເຮັດວຽກ. ຜ່ານປະຫວັດສາດອັນຍາວນານເຮັດໃຫ້ເກີດມີ “ຂໍ້ມູນຈຳນວນມະຫາສານ” ຈາກແຕ່ລະຂະແໜງການລວມທັງພາກລັດ ແລະ ເອກະຊົນທີ່ສາມາດໃຊ້ປະໂຫຍດໄດ້ໂດຍການສະກັດເອົາ ຄວາມຮູ້ໄປຊ່ວຍໃນການພະຍາກອນເຫດການໃດໜຶ່ງ ຫຼື ຊ່ວຍໃນການຕັດສິນໃຈໄດ້ນັ້ນເອງ.

ໃນການຂຽນບົດໂຄງການຈົບຊັ້ນຂອງພວກຂ້າພະເຈົ້າໃນຄັ້ງນີ້ ແມ່ນຂຽນກ່ຽວກັບ ປັນຍາປະດິດຂັ້ນສູງໂດຍມີຫົວຂໍ້ວ່າ: “ການຮຽນຮູ້ ແລະ ການຈື່ຈຳ ຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ ດ້ວຍເຕັກນິກຂອງປັນຍາປະດິດຂັ້ນສູງ” ຈຸດປະສົງໃນການຂຽນບົດ ແມ່ນສຶກສາ, ຮຽນຮູ້ ແລະ ນຳໃຊ້ປັນຍາປະດິດຂັ້ນສູງ ເພື່ອພັດທະນາແບບຈຳລອງການອ່ານຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງດ້ວຍຄວາມສາມາດຂອງພວກນ້ອງເອງ.

ສະນັ້ນ, ພວກຂ້າພະເຈົ້າຈຶ່ງມີຄວາມພາກພູມໃຈຢ່າງຍິ່ງທີ່ໄດ້ສ້າງບົດໂຄງການຈົບຊັ້ນ ສະບັບນີ້ຈົນສຳເລັດ, ອາດມີຂໍ້ບົກຜ່ອງທາງດ້ານເນື້ອໃນບາງປະການທີ່ອະທິບາຍບໍ່ທັນຈະແຈ້ງ ພວກຂ້າພະເຈົ້າກໍຍິນດີທີ່ຈະຮັບຄຳຕິຊົມ ແລະ ຄວາມຄິດເຫັນຈາກບັນດາທ່ານຜູ້ອ່ານ ຫຼື ຜູ້ທີ່ ສົນໃຈທຸກທ່ານເພື່ອຈະປະກອບເປັນບົດຮຽນໃຫ້ພວກຂ້າພະເຈົ້າໃຊ້ໃນການພັດທະນາປັບປຸງໃນຄັ້ງຕໍ່ໄປໃຫ້ສົມບູນຍິ່ງຂຶ້ນໄປອີກ. ພວກຂ້າພະເຈົ້າຫວັງຢ່າງຍິ່ງວ່າບົດໂຄງການຈົບຊັ້ນສະບັບນີ້ຈະເປັນປະໂຫຍດແກ່ຜູ້ທີ່ມີຄວາມສົນໃຈ ຫຼື ມີແນວຄິດຢາກຈະພັດທະນາສິ່ງໃໝ່ດ້ວຍການນຳໃຊ້ເຕັກນິກປັນຍາປະດິດຂັ້ນສູງ (Deep Learning Technology) ເພື່ອພັດທະນາພາສາລາວເຮົາໃຫ້ມີຄົນຮູ້ຈັກຫຼາຍຂຶ້ນ.

ບົດຄັດຫຍໍ້

ການແປຂໍ້ຄວາມຂຽນໄປເປັນຂໍ້ຄວາມສຽງ (Text-To-Speech) ເປັນວິທີການປ່ຽນຂໍ້ຄວາມຂຽນ ຫຼື Text ໄປເປັນຂໍ້ຄວາມສຽງ (Streaming) ເຊິ່ງເປັນນະວັດຕະກໍາໃໝ່ທີ່ມີນັກພັດທະນາຫຼາຍຄົນມີຄວາມສົນໃຈໃນການສຶກສາຄົ້ນຄວ້າ. ການສຶກສາຄົ້ນຄວ້າບົດໂຄງການຈົບຊັ້ນໃນຄັ້ງນີ້ ເປັນການສຶກສາຕົວອັກສອນພາສາລາວ, ສຽງເວົ້າພາສາລາວ ແລະ ໄດ້ປັບປຸງ ພັດທະນາເຄືອຂ່າຍປະສາດທຽມ DNN ຂອງ Deep Learning ໂດຍນໍາໃຊ້ສະຖາປັດຕະຍະກໍາ RNN ດ້ວຍເຕັກນິກ YourTTS ໂດຍການນໍາໃຊ້ວິທີການຂອງການປະມວນຜົນສຽງ. ໃນການເຝິກສອນເຄືອຂ່າຍປະສາດທຽມໃຫ້ຈື່ຈຳສຽງເວົ້າພາສາລາວ ແມ່ນປະກອບໄປດ້ວຍຕົວພະຍັນຊະນະ, ສະຫຼະ, ວັນນະຍຸດ, ສັນຍາລັກອື່ນໆ ແລະ ຂໍ້ຄວາມສຽງເວົ້າພາສາລາວ. ເຊິ່ງລວມຈຳນວນຕົວອັກສອນພາສາລາວທັງໝົດມີ 53 ຕົວ ແລະ ຂໍ້ຄວາມສຽງພາສາລາວ ທີ່ເປັນຟາຍສຽງລວມຈໍານວນທັງໝົດມີ 900 ຟາຍສຽງ ໃນນັ້ນໄດ້ແບ່ງຟາຍສຽງອອກເປັນ 3 ຊຸດຄື: 100% ແມ່ນຊຸດຂໍ້ມູນສຽງທີ່ກຽມໄວ້ເພື່ອໃຫ້ແບບຈໍາລອງຈື່ຈໍາຂໍ້ຄວາມຂຽນ ແລະ ສຽງພາສາລາວ, 95% ແມ່ນຊຸດຂໍ້ມູນສຽງທີ່ກຽມໄວ້ເພື່ອເຝິກສອນແບບຈໍາລອງ ແລະ 5% ແມ່ນຊຸດຂໍ້ມູນສຽງທີ່ກຽມໄວ້ເພື່ອທົດສອບແບບຈຳລອງ.

ໃນຂັ້ນຕອນການຈື່ຈຳສຽງເວົ້າພາສາລາວແມ່ນໄດ້ໃຊ້ແບບຈຳລອງ YourTTS ໃນການປະມວນຜົນ ແລະ ເຝິກສອນແບບຈຳລອງ. ຜົນຂອງການທົດລອງແມ່ນໄດ້ວັດອອກມາເປັນຄ່າ Loss ເຊິ່ງສາມາດວັດປະສິດທິພາບຂອງແບບຈໍາລອງ YourTTS ແມ່ນມີຄ່າການເຝິກສອນໄດ້ຄິດເປັນຄ່າຕ່າງໆຄື: ເວລາສະເລ່ຍໃນການໂຫຼດຂໍ້ມູນ (avg\_loader\_time) ມີຄ່າ 0.0054 ວິນາທີ, ການສູນເສຍ mel-spectrogram (avg\_loss\_mel) ມີຄ່າ 20.29, ເວລາການສູນເສຍການຄາດຄະເນ ແລະ ໄລຍະເວລາຄວາມຈິງພື້ນຖານຂອງ phoneme ໃນການເວົ້າທີ່ສ້າງຂຶ້ນ (avg\_loss\_duration) ມີຄ່າ 0.34 ວິນາທີ ແລະ ການສູນເສຍ mel-spectrogram (avg\_loss\_1) ມີຄ່າ 30.27 ໃນການເຝິກສອນໃນຮອບທີ່ 100,000 ແລະ ໄດ້ນໍາເອົາແບບຈໍາລອງ YourTTS ມາພັດທະນາເປັນ Website ການແປຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ ແລະ ການແປຂໍ້ຄວາມສຽງໄປເປັນຂໍ້ຄວາມຂຽນ

Abstract

Text-To-Speech is a method of converting text to audio (Streaming) which is a new innovation that many developers are interested in researching. The research study of this graduate project is the study of Lao characters, Lao speech and has improved the development of DNN artificial neural network of Deep Learning using RNN architecture with YourTTS technique by using the method of voice processing. In training the artificial neural network to remember the speech in Lao language is composed of consonants, vowels, tones, other symbols and voice messages in Lao language. which includes 53 Lao alphabets and Lao audio text that is a total of 900 audio files, in which the audio file is divided into 3 sets: 100% is a set of audio data prepared for the model to remember written text and Lao language, 95% is a set of audio data prepared to train the model and 5% is a set of audio data prepared to test the model.

In the Lao speech recognition process, the YourTTS model is used in processing and training the model. The result of the experiment is measured as a Loss value which can measure the effectiveness of the YourTTS model. The training value can be calculated as: avg\_loader\_time with a value of 0.0054 seconds, avg\_loss\_mel with a value of 20.29, avg\_loss\_duration with a value of 0.345 seconds and avg\_loss\_1 with a value of 30.27 in the training cycle of 100,000 and brought the model YourTTS is developed as a website to translate written Lao text into voice text and to translate voice text into written text

ຄຳສະແດງຄຳຂອບໃຈ

ບົດໂຄງການຈົບຊັ້ນສະບັບນີ້ ຂອງພວກຂ້າພະເຈົ້າ ທ້າວ ຄຳປະເສີດ ໄຊຍະວົງ ແລະ ທ້າວ ໃຫຍ່ ແສງວິໄຊ ໄດ້ສຳເລັດຢ່າງສົມບູນ ກໍຍ້ອນໄດ້ຮັບຄຳປຶກສາ, ແນະນຳ ແລະ ການຊ່ວຍເຫຼືອ ຈາກບຸກຄົນທີ່ຮັກແພງ ແລະ ເຄົາລົບນັບຖືຫຼາຍໆທ່ານ. ດັ່ງນັ້ນ, ພວກ ຂ້າພະເຈົ້າໃນນາມນັກສຶກສາ ມະຫາວິທະຍາໄລແຫ່ງຊາດ, ສັງກັດຢູ່ຄະນະວິທະຍາສາດທຳມະຊາດ, ພາກວິຊາວິທະຍາສາດຄອມພິວເຕີ ປີທີ 4 ລຸ້ນທີ XXIII ສົກສຶກສາ 2022-2023 ຂໍສະແດງຄວາມຮູ້ບຸນຄຸນ ແລະ ເຄົາລົບນັບຖືຕໍ່ກັບຄະນະນຳ, ບັນດາຄູອາຈານທຸກພາກສ່ວນທີ່ກ່ຽວຂ້ອງພາຍໃນ ຄະນະວິທະຍາສາດທຳມະຊາດ ໂດຍສະເພາະແມ່ນພາກວິຊາ ວິທະຍາສາດຄອມພິວເຕີ ທີ່ປຽບດັ່ງແສງສະຫວ່າງທີ່ຊີ້ທິດເຍືອງທາງໃຫ້ແກ່ພວກຂ້າພະເຈົ້າ ທີ່ໄດ້ທຸ້ມເທໃສ່ການອົບຮົມ ແນວຄິດສິດສອນຄວາມຮູ້ຕ່າງໆ, ສ້າງແຮງບັນດານໃຈໃຫ້ແກ່ພວກຂ້າພະເຈົ້າສຳເລັດການຂຽນບົດໂຄງການຈົບຊັ້ນຄັ້ງນີ້ກໍຄື ສຳເລັດການສຶກສາລະດັບປະລິນຍາຕີ.

ພວກຂ້າພະເຈົ້າຂໍຖືໂອກາດທີ່ພາກພູມໃຈນີ້ສະແດງຄໍາຂອບໃຈ ແລະ ຮູ້ບຸນຄຸນຢ່າງສູງ ມາຍັງ ອຈ.ປອ. ລັດສະໝີ ຈິດຕະວົງ ເຊິ່ງເປັນອາຈານຜູ້ນຳພາທີ່ໄດ້ໃຫ້ຄຳແນະນຳ ແລະ ເປັນທີ່ ປືກສາຕະຫຼອດໄລຍະໃນການສຶກສາຮ່ຳຮຽນ ແລະ ການຂຽນບົດໂຄງການຈົບຊັ້ນໃນຄັ້ງນີ້ຈົນໄດ້ ຮັບຜົນສຳເລັດ. ຂອບໃຈ ອຈ.ປທ. ສົມມິດ ທຸມມາລີ ເຊິ່ງເປັນອາຈານຜູ້ຊ່ວຍນຳພາບົດໂຄງການ ຈົບຊັ້ນໃນຄັ້ງນີ້ ໃຫ້ມີເນື້ອໃນຄົບຖ້ວນ ແລະ ສົມບູນ.

ເພີ່ມເຕີມ, ຂໍສະແດງຄວາມຂອບໃຈ ແລະ ຮູ້ບຸນຄຸນມາຍັງອາຈານທຸກທ່ານພາຍໃນ ຄະນະວິທະຍາສາດ ທຳມະຊາດ ກໍຄືພາກວິຊາວິທະຍາສາດຄອມພິວເຕີທີ່ໄດ້ເປັນບ່ອນເຝິກຝົນຫຼໍ່ ຫຼອມ, ມອບຄວາມຮູ້ວິຊາຕ່າງໆ ແລະ ຊີ້ນຳທາງໃນການສຶກສາຮໍ່າຮຽນໃຫ້ແກ່ພວກຂ້າພະເຈົ້າ ຕັ້ງແຕ່ຕົ້ນຈົນຈົບການສຶກສາທີ່ໄດ້ມາສຶກສາໃນສະຖາບັນແຫ່ງນີ້.

ສຸດທ້າຍນີ້, ຂໍສະແດງຄວາມຂອບໃຈມາຍັງພໍ່, ແມ່, ຄອບຄົວ, ອ້າຍເອື້ອຍນ້ອງ ແລະ ໝູ່ເພື່ອນທີ່ສະໜັບສະໜູນ ຊ່ວຍເປັນແຮງກຳລັງໃຈຕະຫຼອດມາ, ຊ່ວຍແກ້ໄຂບັນຫາຈົນມາເຖິງຂັ້ນສຸດທ້າຍຂອງການສຶກສາກໍຄືການຂຽນບົດໂຄງການຈົບຊັ້ນໃນລະດັບປະລິນຍາຕີທີ່ສະຖາບັນແຫ່ງນີ້ ແລະ ຍັງຊ່ວຍເຫຼືອທາງດ້ານວັດຖຸເງິນຄຳໃນທຸກໆດ້ານ.

ສາລະບານ

ໜ້າ

[ຄຳນຳ i](#_Toc141244731)

[ບົດຄັດຫຍໍ້ ii](#_Toc141244732)

[Abstract iii](#_Toc141244733)

[ຄຳສະແດງຄຳຂອບໃຈ iv](#_Toc141244734)

[ສາລະບານ v](#_Toc141244735)

[ສາລະບານ (ຕໍ່) v](#_Toc141244735)i

[ສາລະບານ (ຕໍ່) v](#_Toc141244735)ii

[ສາລະບານຕາຕະລາງ viii](#_Toc141244736)

[ສາລະບານແຜນວາດ ix](#_Toc141244737)

[ສາລະບານຮູບພາບ x](#_Toc141244738)

[ສາລະບານຮູບພາບ (ຕໍ່) xi](#_Toc141244739)

[ສາລະບານຮູບພາບ (ຕໍ່) xi](#_Toc141244739)i

[ບົດທີ 1 ບົດສະເໜີ 1](#_Toc141244740)

[1.1 ຄວາມສຳຄັນຂອງບັນຫາ 1](#_Toc141244741)

[1.2 ຈຸດປະສົງຂອງການຄົ້ນຄ້ວາ (Objectives) 2](#_Toc141244742)

[1.3 ຂອບເຂດຂອງການຄົ້ນຄ້ວາ (Scope) 2](#_Toc141244743)

[1.4 ປະໂຫຍດຄາດວ່າຈະໄດ້ຮັບ (Expected Outcome of the Project) 3](#_Toc141244744)

[1.5 ອະທິບາຍຄຳສັບ 3](#_Toc141244745)

[ບົດທີ 2 ທົບທວນເອກະສານ ແລະ ບົດຄົ້ນຄວ້າທີ່ກ່ຽວຂ້ອງ 5](#_Toc141244746)

[2.1 ທົບທວນເອກະສານທີ່ກ່ຽວຂ້ອງ 5](#_Toc141244748)

[2.1.1 ການປະມວນຜົນພາສາທຳມະຊາດ (Natural Language Processing) 5](#_Toc141244758)

[2.1.2 ເຕັກນິກການຮຽນຮູ້ຂັ້ນສູງ (Deep Learning) 6](#_Toc141244759)

[2.1.3 ເຕັກນິກຂອງ YourTTS 12](#_Toc141244760)

[2.1.4 Google Speech-To-Text API 13](#_Toc141244761)

[2.1.5 Unified Modeling Language (UML) 14](#_Toc141244762)

[2.1.6 ຄວາມຮູ້ກ່ຽວກັບ Flowchart 20](#_Toc141244763)

[2.2 ບົດຄົ້ນຄວ້າທີ່ກ່ຽວຂ້ອງ 21](#_Toc141244764)

[ບົດທີ 3 ວິທີດໍາເນີນການຄົ້ນຄວ້າ 24](#_Toc141244765)

ສາລະບານ (ຕໍ່)

ໜ້າ

[3.1 ວິທີສຶກສາ ແລະ ຄົ້ນຄວ້າ 24](#_Toc141244767)

[3.1.1 ການກຳນົດເນື້ອໃນ 24](#_Toc141244772)

[3.1.2 ການຄັດເລືອກພື້ນທີ່ 25](#_Toc141244773)

[3.2 ເຄື່ອງມືທີ່ໃຊ້ໃນການຄົ້ນຄວ້າ 26](#_Toc141244774)

[3.3 ການເກັບຮວບຮວມຂໍ້ມູນ 27](#_Toc141244779)

[3.3.1 ການຄັດເລືອກປະຊາກອນ 27](#_Toc141244794)

[3.3.2 ການກຳນົດຈຳນວນຕົວຢ່າງປະຊາກອນ 29](#_Toc141244797)

[3.3.3 ວິທີສຸ່ມຕົວຢ່າງປະຊາກອນ 30](#_Toc141244798)

[3.3.4 ບັນດາຂໍ້ມູນການສຶກສາ 31](#_Toc141244799)

[3.3.5 ວິທີເກັບກຳຂໍ້ມູນ 32](#_Toc141244800)

[3.4 ການວິເຄາະຂໍ້ມູນ 33](#_Toc141244801)

[3.4.1 ຂະບວນການວິເຄາະຂໍ້ມູນ 33](#_Toc141244821)

[3.4.2 ການນຳແບບຈຳລອງໄປໃຊ້ງານ 36](#_Toc141244822)

[3.5 ການວິເຄາະລະບົບ 36](#_Toc141244823)

[3.5.1 ແຜນວາດລວມຂອງລະບົບ 36](#_Toc141244828)

[3.5.2 ແຜນວາດ Use Case Diagram 37](#_Toc141244829)

3.5.3 ແຜນວາດ Sequence Diagram 38

3.5.4 ແຜນວາດ Activity Diagram 40

[3.6 ການອອກແບບລະບົບ 41](#_Toc141244830)

[3.6.1 ການອອກແບບຮ່າງສະແດງຜົນ 41](#_Toc141244855)

[3.6.2 ການອອກແບບໂປຣແກຣມ 43](#_Toc141244856)

ບົດທີ 4 ຜົນຂອງການສຶກສາ ແລະ ການອະທິບາຍຜົນ 45

[4.1 ການລາຍງານຜົນຂອງການຄົ້ນຄວ້າ 45](#_Toc141244859)

[4.1.1 ລາຍງານຜົນການສຶກສາຂັ້ນຕອນການເຝິກສອນແບບຈຳລອງ 46](#_Toc141244871)

[4.1.2 ລາຍງານຜົນການສຶກສາຂັ້ນຕອນການທົດລອງແບບຈຳລອງ 49](#_Toc141244872)

[4.1.3 ລາຍງານຜົນການສຶກສາການປະເມີນແບບຈຳລອງ 52](#_Toc141244873)

[4.1.4 ລາຍງານຜົນການຂຽນເວັບໄຊ ແລະ ວິທີການນຳໃຊ້ 54](#_Toc141244874)

[4.2 ການອະທິບາຍຜົນຂອງການຄົ້ນຄວ້າ 56](#_Toc141244875)

[ບົດທີ 5 ສະຫຼຸບ ແລະ ຂໍ້ສະເໜີ 58](#_Toc141244876)

[5.1 ສະຫຼຸບການຄົ້ນຄວ້າ 58](#_Toc141244878)

ສາລະບານ (ຕໍ່)

ໜ້າ

[5.2 ຂໍ້ຈຳກັດຂອງການຄົ້ນຄວ້າ 59](#_Toc141244879)

[5.3 ຂໍ້ສະເໜີໃນການຄົ້ນຄວ້າຕໍ່ໄປ 59](#_Toc141244880)

[ເອກະສານອ້າງອີງ 60](file:///C:\Users\gnay17\Documents\Yai_2023\Proposal\Proposal%20Final\Text_To_Speech(250723).docx#_Toc141244881)

[ເອກະສານຊ້ອນທ້າຍ 65](file:///C:\Users\gnay17\Documents\Yai_2023\Proposal\Proposal%20Final\Text_To_Speech(250723).docx#_Toc141244882)

[ປະຫວັດຫຍໍ້ຜູ້ຂຽນບົດ 67](file:///C:\Users\gnay17\Documents\Yai_2023\Proposal\Proposal%20Final\Text_To_Speech(250723).docx#_Toc141244883)

ສາລະບານຕາຕະລາງ

ໜ້າ

[ຕາຕະລາງທີ 2.1 ຕາຕະລາງສະແດງສັນຍາລັກທີ່ໃຊ້ແຕ້ມແຜນວາດ Flowchart 21](#_Toc141243427)

[ຕາຕະລາງທີ 3.1 ຕົວອັກສອນໃນພາສາລາວ 27](#_Toc141243428)

[ຕາຕະລາງທີ 3.2 ບົດຄວາມ ແລະ ສື່ສິ່ງພິມຕ່າງໆທາງອິນເຕີເນັດ 28](#_Toc141243429)

ຕາຕະລາງທີ 3.3 ຄໍາອະທິບາຍ Text-To-Speech 38

ຕາຕະລາງທີ 3.4 ຄໍາອະທິບາຍ Speech-To-Text 38

ສາລະບານແຜນວາດ

ໜ້າ

[ແຜນວາດທີ 3.1 ແຜນວາດສະແດງການໃຊ້ງານ Website 36](#_Toc141246832)

[ແຜນວາດທີ 3.2 ແຜນວາດສະແດງ Use Case Diagram ຂອງ Website ການແປຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ ແລະ ການແປຂໍ້ຄວາມສຽງໄປເປັນຂໍ້ຄວາມຂຽນ 37](#_Toc141246833)

ແຜນວາດທີ 3.3 ແຜນວາດສະແດງ Sequence Diagram ຂອງ Text-To-Speech 39

ແຜນວາດທີ 3.4 ແຜນວາດສະແດງ Sequence Diagram ຂອງ Speech-To-Text 39

ແຜນວາດທີ 3.5 ແຜນວາດສະແດງ Activiti Diagram ຂອງ Text-To-Speech 40

ແຜນວາດທີ 3.6 ແຜນວາດສະແດງ Activiti Diagram ຂອງ Speech-To-Text 41

[ແຜນວາດທີ 3.7 ແຜນວາດ Flowchart ສະແດງເຖິງການເຮັດວຽກຂອງ Text-To-Speech 43](#_Toc141246834)

ແຜນວາດທີ 3.8 ແຜນວາດ Flowchart ສະແດງເຖິງການເຮັດວຽກຂອງ Speech-To-Text 44

ສາລະບານຮູບພາບ

ໜ້າ

[ຮູບທີ 2.1 ໂຄງສ້າງໂຄງຂ່າຍປະສາດທຽມໂດຍມີສ່ວນຍ່ອຍໃນການຄໍານວນຄື Node ແລະ ການລຽງຕົວຂອງ Node ເປັນຊັ້ນ ເອີ້ນວ່າ: Layer 7](#_Toc141244118)

[ຮູບທີ 2.2 ຕົວຢ່າງໂຄງສ້າງຂອງຂັ້ນຕອນວິທີ Deep Neural Network 9](#_Toc141244119)

[ຮູບທີ 2.3 ຕົວຢ່າງການເຮັດວຽກໃນເຊວຂອງຂັ້ນຕອນວິທີໂຄງຂ່າຍປະສາດທຽມ RNN 10](#_Toc141244120)

[ຮູບທີ 2.4 ແຜນວາດການເຝິກສອນແບບຈໍາລອງ ແລະ ອະນຸມານແບບຈໍາລອງ YourTTS 13](#_Toc141244121)

[ຮູບທີ 2.5 ຮູບພາບສະແດງອົງປະກອບຂອງ UML 15](#_Toc141244122)

[ຮູບທີ 2.6 ຮູບພາບສະແດງສັນຍາລັກ Use Case 15](#_Toc141244123)

[ຮູບທີ 2.7 ຮູບພາບສະແດງສັນຍາລັກ Actor 16](#_Toc141244124)

[ຮູບທີ 2.8 ຮູບພາບສະແດງສັນຍາລັກ System Boundary 16](#_Toc141244125)

[ຮູບທີ 2.9 ຮູບພາບສະແດງຄວາມສໍາພັນ Include Relationship 16](#_Toc141244126)

[ຮູບທີ 2.10 ຮູບພາບສະແດງຄວາມສໍາພັນ Extend Relationship 17](#_Toc141244127)

[ຮູບທີ 2.11 ຮູບພາບສະແດງຄວາມສໍາພັນ Generalization 17](#_Toc141244128)

ຮູບທີ 2.12 ຮູບພາບສະແດງສັນຍາລັກຂອງ Actor ໃນ Sequence Diagram 18

ຮູບທີ 2.13 ຮູບພາບສະແດງສັນຍາລັກຂອງ Object ໃນ Sequence Diagram 18

ຮູບທີ 2.14 ຮູບພາບສະແດງສັນຍາລັກຂອງ Lifeline ໃນ Sequence Diagram 18

ຮູບທີ 2.15 ຮູບພາບສະແດງສັນຍາລັກຂອງ Activation ໃນ Sequence Diagram 18

ຮູບທີ 2.16 ຮູບພາບສະແດງສັນຍາລັກຂອງ Message ໃນ Sequence Diagram 19

ຮູບທີ 2.17 ຮູບພາບສະແດງສັນຍາລັກຂອງ Callback ໃນ Sequence Diagram 19

ຮູບທີ 2.18 ຮູບພາບສະແດງ Sequence Diagram 19

ຮູບທີ 2.19 ຮູບພາບສະແດງ Activity Diagram 20

[ຮູບທີ 3.1 ຮູບພາບສະແດງພາບລວມຂອງເຕັກນິກເຄືອຂ່າຍປະສາດທຽມ DNN ຂອງ Deep Learning ໂດຍນໍາໃຊ້ສະຖາປັດຕະຍະກໍາ RNN ດ້ວຍເຕັກນິກ YourTTS 24](#_Toc141244129)

[ຮູບທີ 3.2 ຂັ້ນຕອນການສຶກສາ 25](#_Toc141244130)

[ຮູບທີ 3.3 ຮູບພາບສະແດງຂະບວນການເກັບກຳຊຸດຂໍ້ມູນໃນ Google Drive 30](#_Toc141244131)

[ຮູບທີ 3.4 ຮູບພາບສະແດງໃຫ້ເຫັນເຖິງການແບ່ງຂໍ້ມູນເພື່ອການ Train ແລະ Test 31](#_Toc141244132)

[ຮູບທີ 3.5 ຮູບພາບສະແດງລະດັບຂອງຕົວອັກສອນ 32](#_Toc141244133)

ສາລະບານຮູບພາບ (ຕໍ່)

ໜ້າ

[ຮູບທີ 3.6 ຮູບພາບສະແດງຊຸດຂໍ້ມູນທັງໝົດໃນຟາຍ metadata.csv 33](#_Toc141244134)

[ຮູບທີ 3.7 ຮູບພາບສະແດງຂັ້ນຕອນການປະມວນຜົນແບບຈຳລອງ YourTTS 34](#_Toc141244135)

[ຮູບທີ 3.8 ຮູບພາບສະແດງການອອກແບບພາກສ່ວນການແປຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ 42](#_Toc141244136)

[ຮູບທີ 3.9 ຮູບພາບສະແດງການອອກແບບພາກສ່ວນການແປຂໍ້ຄວາມສຽງພາສາລາວໄປເປັນຂໍ້ຄວາມຂຽນ 42](#_Toc141244137)

ຮູບທີ 4.1 ຮູຮູບພາບສະແດງຄ່າຕ່າງໆໃນການປະມວນຜົນແບບຈໍາລອງໃນຮອບທີ 100,000 45

[ຮູບທີ 4.2 ຮູບພາບສະແດງເວລາສະເລ່ຍໃນການໂຫຼດຂໍ້ມູນການໃນລະຫວ່າງການເຝິກສອນແບບຈຳລອງ (TrainIterStats/loader\_time) 46](#_Toc141244139)

[ຮູບທີ 4.3 ຮູບພາບສະແດງໄລຍະເວລາການສູນເສຍການຄາດຄະເນແລະໄລຍະເວລາຄວາມຈິງພື້ນຖານຂອງ phoneme ໃນການເວົ້າທີ່ສ້າງຂຶ້ນໃນລະຫວ່າງການເຝິກສອນແບບຈຳລອງ (TrainIterStats/loss\_duration) 47](#_Toc141244140)

[ຮູບທີ 4.4 ຮູບພາບສະແດງການສູນເສຍຄັ້ງທໍາອິດທີ່ໃຊ້ໃນລະຫວ່າງການເຝິກສອນຂອງ vocoder MelGAN (TrainIterStats/loss\_01) 48](#_Toc141244141)

[ຮູບທີ 4.5 ຮູບພາບສະແດງການສູນເສຍ mel-spectrogram ໃນລະຫວ່າງການເຝິກສອນແບບຈຳລອງ (TrainIterStats/loss\_mel) 49](#_Toc141244142)

[ຮູບທີ 4.6 ຮູບພາບສະແດງເວລາສະເລ່ຍໃນການໂຫຼດຂໍ້ມູນໃນລະຫວ່າງການທົດລອງ (EvalStats/avg\_loader\_time) 50](#_Toc141244143)

[ຮູບທີ 4.7 ຮູບພາບສະແດງການສູນເສຍ mel-spectrogram (EvalStats/avg\_loss\_mel) 50](#_Toc141244144)

[ຮູບທີ 4.8 ຮູບພາບສະແດງໄລຍະເວລາການສູນເສຍການຄາດຄະເນແລະໄລຍະເວລາຄວາມຈິງພື້ນຖານຂອງ phoneme ໃນການເວົ້າທີ່ສ້າງຂຶ້ນ (EvalStats/ avg\_loss\_ duration) 51](#_Toc141244145)

[ຮູບທີ 4.9 ຮູບພາບສະແດງການສູນເສຍ mel-spectrogram ໃນລະຫວ່າງການທົດລອງ (EvalStats/ avg\_loss\_ 1) 52](#_Toc141244146)

[ຮູບທີ 4.10 ຮູບພາບສະແດງຕົວຊີ້ວັດທີ່ໃຊ້ໃນແບບຈຳລອງ ເພື່ອປະເມີນຄຸນນະພາບຂອງການຈັດຮຽງລະຫວ່າງຂໍ້ຄວາມທີ່ປ້ອນເຂົ້າ ແລະ ຮູບແບບຄື້ນສຽງທີ່ສ້າງຂຶ້ນ 53](#_Toc141244147)

[ຮູບທີ 4.11 ຮູບພາບສະແດງຕົວຊີ້ວັດທີ່ໃຊ້ ເພື່ອປຽບທຽບຄຸນນະພາບຂອງສຽງເວົ້າທີ່ສັງເຄາະດ້ວຍສັນຍານສຽງເວົ້າຕົ້ນສະບັບ 54](#_Toc141244148)

[ຮູບທີ 4.12 ຮູບພາບສະແດງຂັ້ນຕອນວິທີການນໍາໃຊ້ເວັບໄຊໃນພາກສ່ວນການແປຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ 55](#_Toc141244149)

ສາລະບານຮູບພາບ (ຕໍ່)

ໜ້າ

[ຮູບທີ 4.13 ຮູບພາບສະແດງຂັ້ນຕອນວິທີການນໍາໃຊ້ເວັບໄຊໃນພາກສ່ວນການແປຂໍ້ຄວາມສຽງພາສາລາວໄປເປັນຂໍ້ຄວາມຂຽນ 56](#_Toc141244150)

# ບົດສະເໜີ

## ຄວາມສຳຄັນຂອງບັນຫາ

ຊາດລາວ ເປັນຊາດໜຶ່ງທີ່ມີພາສາ ແລະ ຕົວອັກສອນເປັນຂອງຕົນເອງມີຊື່ຮຽກວ່າ ພາສາລາວ ແລະ ຕົວອັກສອນລາວ. ສໍາລັບພາສາລາວນັ້ນມັນກຳເນີດຂຶ້ນພ້ອມໆ ກັບການກໍາເນີດຂອງຊາດລາວ ເປັນພາສາທີ່ເກົ່າແກ່ພາສາໜຶ່ງຢູ່ໃນອາຊີອາຄະເນ. ສ່ວນຕົວອັກສອນລາວມີໃຊ້ແລ້ວໃນສະໄໝຂຸນຫລວງສີເມົາ ເຈົ້າປົກຄອງເມືອງງາຍລາວ. ທັງພາສາ ແລະ ຕົວອັກສອນລາວ ໄດ້ຮັບການພັດທະນາເລື້ອຍໆ. ໃນອະດີດອັນຍາວນານຊົນຊາດລາວມີອໍານາດ, ມີອິດທິພົນແຜ່ຂະຫຍາຍກວ້າງໄກ, ພາສາ ແລະ ຕົວອັກສອນລາວກໍ່ໄດ້ຮັບການພັດທະນາໃຫ້ດີຂຶ້ນເພື່ອຮັບໃຊ້ໃຫ້ແກ່ການ ຂະຫຍາຍຕົວທາງດ້ານເສດຖະກິດ, ວັດທະນະທຳສັງຄົມຢ່າງກວ້າງຂວາງ (ກະຊວງໄປສະນີ ໂທລະຄົມມະນາຄົມ ແລະ ການສື່ສານ, 2012).

ໃນປີ 2003, ອົງການວິທະຍາສາດເຕັກໂນໂລຊີ ແລະ ສິ່ງແວດລ້ອມຂອງລາວ ໂດຍໄດ້ຮັບການຊ່ວຍເຫຼືອ ຈາກອົງການສະຫະປະຊາຊາດເພື່ອການພັດທະນາໄດ້ຈັດຕັ້ງໂຄງການສ້າງມາດຕະຖານຕົວອັກສອນລາວໃນລະບົບຄອມພິວເຕີ ໂດຍມີ ທ່ານ ສະນິດ ຊາພັກດີ ເປັນຫົວໜ້າຊີ້ນຳ ແລະ ທ່ານ ອານຸສັກ ສຸພາວັນ ເປັນ ຫົວໜ້າທີມງານນັກຄົ້ນຄ້ວາ ແລະ ໄດ້ສ້າງ ຟອນເພັດຊະລາດ ແລະ ຊອບແວລະບົບເປີດຕ່າງໆ ເປັນພາສາລາວ ເຊັ່ນ Lao Linux, Lao Open Office ແລະ Lao Thunderbird. ເພື່ອເຮັດໃຫ້ມີຄວາມເປັນເອກກະພາບໃນການນໍາໃຊ້ຕົວອັກສອນລາວໃນລະບົບຄອມພິວເຕີໃນເດືອນ ກໍລະກົດ 2009 ລັດຖະບານໄດ້ປະກາດຮັບຮອງເອົາຟອນເພັດຊະລາດ (Phetsarath OT) ເປັນມາດຕະຖານແຫ່ງຊາດຂອງການນໍາໃຊ້ຕົວອັກສອນລາວໃນລະບົບຄອມພິວເຕີ (ກະຊວງໄປສະນີ ໂທລະຄົມມະນາຄົມ ແລະ ການສື່ສານ, 2012).

ໃນປັດຈຸບັນໂລກຂອງພວກເຮົາໄດ້ມີການພັດທະນາຄວາມສາມາດຂອງອຸປະກອນຕ່າງໆ ດ້ານຄອມພິວເຕີ ໃຫ້ມີການຮຽນຮູ້ຕາມແບບຢ່າງຂອງມະນຸດ. ໃນດ້ານການຈື່ຈຳສຽງເວົ້າ (Speech Recognition) ກໍມີຄວາມກ້າວໜ້າໄປຫຼາຍ ໄດ້ມີການນຳໄປປະຍຸກໃຊ້ໃນລະບົບສື່ສານໂທລະຄົມມະນາຄົມ ເຊັ່ນ: ການສັ່ງໂທອອກຂອງໂທລະສັບມືຖືໂດຍໃຊ້ສຽງ. ຈຸດປະສົງຫຼັກຂອງການຈື່ຈຳສຽງເວົ້າຄືການເພີ່ມຄວາມສາມາດໃຫ້ອຸປະກອນຕ່າງໆ ສາມາດຮັບຮູ້ ແລະ ມີການໂຕ້ຕອບກັບມະນຸດໄດ້ຫຼາຍຂຶ້ນ ເຊິ່ງການໃຊ້ສຽງເວົ້າໃນການສື່ສານຄວບຄຸມສັ່ງການຖືວ່າເປັນວິທີທີ່ມະນຸດໃຊ້ຢ່າງເປັນທຳມະຊາດທີ່ສຸດ. ເນື່ອງຈາກວ່າ, ພາສາລາວຍັງບໍ່ມີນັກຄົ້ນຄວ້າທ່ານໃດທີ່ສ້າງລະບົບການແປຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ (Lao Text To Speech) ມາເຜີຍແຜ່ ແລະ ໃຫ້ນຳໃຊ້ກັນໃນດ້ານອຸປະກອນຄອມພິວເຕີຕ່າງໆ. ສະນັ້ນ, ພວກຂ້າພະເຈົ້າ ຈຶ່ງມີຄວາມສົນໃຈຢາກສຶກສາ, ປັບປຸງ ແລະ ນຳໃຊ້ແບບຈຳລອງການແປຂໍ້ຄວາມຂຽນໄປເປັນຂໍ້ຄວາມສຽງ (Text-To-Speech) ຂອງພາສາລາວໃຫ້ມີປະສິດທິພາບ, ສາມາດຮຽນຮູ້ ແລະ ຈື່ຈຳ ໄດ້ຖືກຕ້ອງຊັດເຈນຍິ່ງຂຶ້ນ.

## ຈຸດປະສົງຂອງການຄົ້ນຄ້ວາ (Objectives)

* ເພື່ອສຶກສາຂະບວນການເຮັດວຽກຂອງ ລະບົບການຈື່ຈຳການແປຂໍ້ຄວາມຂຽນໄປເປັນຂໍ້ຄວາມສຽງ (Text-To-Speech) ໃນການຮຽນຮູ້, ຈື່ຈຳ ພາສາລາວ ແລະ ເພື່ອສຶກສາການປະມວນຜົນ ແລະ ວິເຄາະພາສາລາວມີທັງຂໍ້ຄວາມຂຽນ ແລະ ຂໍ້ຄວາມສຽງ ໂດຍການພັດທະນາມາຈາກເຄືອຂ່າຍປະສາດທຽມ DNN ຂອງ Deep Learning ໂດຍນໍາໃຊ້ສະຖາປັດຕະຍະກໍາ RNN ດ້ວຍເຕັກນິກ YourTTS.
* ເພື່ອປະມວນຜົນການຮຽນຮູ້ ແລະ ການຈື່ຈຳຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ (Lao Text-To-Speech) ດ້ວຍແບບຈໍາລອງ YourTTS ທີ່ປັບປຸງແລ້ວ.
* ເພື່ອສະແດງຜົນການຮຽນຮູ້ ແລະ ການຈື່ຈຳຂໍ້ຄວາມສຽງພາສາລາວໄປເປັນຂໍ້ຄວາມຂຽນ (Lao Speech-To-Text) ໂດຍນໍາໃຊ້ແບບຈໍາລອງ Google Speech To Text API.
* ເພື່ອສ້າງເວັບໄຊການແປຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ ແລະ ການແປຂໍ້ຄວາມສຽງພາສາລາວໄປເປັນຂໍ້ຄວາມຂຽນ ທີ່ມີປະສິດທິພາບ.

## ຂອບເຂດຂອງການຄົ້ນຄ້ວາ (Scope)

* ກະກຽມຂໍ້ມູນສຽງພາສາລາວສຳນຽງພາກກາງ ເພື່ອນຳໄປໃຊ້ໃນການຮຽນຮູ້ ແລະ ຈື່ຈຳ ຂໍ້ຄວາມຂຽນພາສາລາວ ເປັນຈຳນວນ 600 ຄໍາ.
* ສ້າງຂັ້ນຕອນວິທີແບບ DNN ໂດຍນຳໃຊ້ວິທີການຮຽນຮູ້ແບບການແປຂໍ້ຄວາມຂຽນໄປເປັນຂໍ້ຄວາມສຽງ (Text-To-Speech) ໂດຍນໍາໃຊ້ສະຖາປັດຕະຍະກໍາ RNN ດ້ວຍເຕັກນິກ YourTTS ເພື່ອສ້າງແບບຈຳລອງການແປຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ (Lao Text-To-Speech).
* ປະເມີນປະສິດທິພາບ ແລະ ການທົດສອບແບບຈຳລອງການແປຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ (Lao Text-To-Speech).
* ສຶກສາ ແລະ ນຳໃຊ້ແບບຈຳລອງຂໍ້ມູນທີ່ຖືກພັດທະນາຂຶ້ນໂດຍ Google Speech To Text API ເພື່ອແປຄຳເວົ້າພາສາລາວໄປເປັນຂໍ້ຄວາມຂຽນ (Lao Speech To Text).
* ສ້າງເວັບໄຊການແປຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ ແລະ ການແປຂໍ້ຄວາມສຽງໄປເປັນຂໍ້ຄວາມຂຽນ.

## ປະໂຫຍດຄາດວ່າຈະໄດ້ຮັບ (Expected Outcome of the Project)

* ໄດ້ຮູ້ເຖິງຂະບວນການເຮັດວຽກ ແລະ ບັນຫາທີ່ເກີດຂຶ້ນຂອງ ແບບຈໍາລອງການແປຂໍ້ຄວາມຂຽນໄປເປັນຂໍ້ຄວາມສຽງ (Text-To-Speech).
* ໄດ້ແບບຈຳລອງການແປຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ (Lao Speech-To-Text) ໃໝ່ ທີ່ມີປະສິດທິພາບ.
* ໄດ້ເວັບໄຊການແປຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ ແລະ ການແປຂໍ້ຄວາມສຽງພາສາລາວໄປເປັນຂໍ້ຄວາມຂຽນ ທີ່ມີປະສິດທິພາບ.

## ອະທິບາຍຄຳສັບ

###### VITS (ຕົວເຂົ້າລະຫັດອັດຕະໂນມັດແບບຜັນປ່ຽນຕາມເງື່ອນໄຂພ້ອມ Adversarial Learning ສຳລັບການອ່ານອອກສຽງຂໍ້ຄວາມແບບ End-to-End) ເປັນແບບຈຳລອງ TTS ແບບ End-to-End (ຕົວເຂົ້າລະຫັດ −> vocoder ຮ່ວມກັນ) ທີ່ໃຊ້ປະໂຫຍດຈາກເຕັກນິກ SOTA DL ເຊັ່ນ GANs, VAE ແລະ Normalizing Flows (Kim, Kong and Son, 2021).

###### TTS ຫຍໍ້ມາຈາກ Text-To-Speech ໝາຍເຖິງການແປຂໍ້ຄວາມຂຽນໄປເປັນຂໍ້ຄວາມສຽງ

###### STT ຫຍໍ້ມາຈາກ Speech-To-Text ໝາຍເຖິງການແປຂໍ້ຄວາມສຽງໄປເປັນຂໍ້ຄວາມຂຽນ

###### DNN ຫຍໍ້ມາຈາກ Deep Neural Network ແມ່ນ Algorithm ທີ່ຈັດຢູ່ໃນ Deep Learning

###### RNN ຫຍໍ້ມາຈາກ Recurrent Neural Network ແມ່ນສະຖາປັດຕະຍະກໍາທີ່ຈັດຢູ່ໃນ Deep Learning

###### YourTTS ແມ່ນແບບຈໍາລອງການແປຂໍ້ຄວາມຂຽນໄປເປັນຂໍ້ຄວາມສຽງ ທີ່ສາມາດເວົ້າໄດ້ຫຼາຍພາສາ ຫຼື ເອີ້ນວ່າ Multilingual

###### mel-spectrogram ແມ່ນເປັນຕົວແທນຂອງຄວາມຫນາແຫນ້ນຂອງພະລັງງານສັນຍານ, ຄິດໄລ່ໂດຍໃຊ້ mel-scale filterbank. ມັນຖືກນໍາໃຊ້ຢ່າງກວ້າງຂວາງໃນການປະມວນຜົນສຽງເວົ້າ ແລະ ການວິເຄາະ. ລວມທັງໃນລະບົບຂໍ້ຄວາມເປັນສຽງເວົ້າ (TTS) ເຊັ່ນ YourTTS.

###### phonemes ແມ່ນຫົວໜ່ວຍສຽງໃດນຶ່ງທີ່ເຫັນໄດ້ຊັດເຈນໃນພາສາໃດໜຶ່ງທີ່ຈຳແນກຄຳໜຶ່ງຈາກອີກຄຳໜຶ່ງ, ຕົວຢ່າງ p, b, d, ແລະ t ໃນຄຳສັບພາສາອັງກິດ pad, pat, bad, ແລະ bat.

###### mel-scale filterbank ແມ່ນການປະຕິບັດທາງຄະນິດສາດທີ່ປ່ຽນ mel-spectrogram ກັບຄືນໄປບ່ອນໂດເມນຄວາມຖີ່ເສັ້ນຊື່. ການປະຕິບັດງານນີ້ແມ່ນມີຄວາມຈໍາເປັນໃນລະບົບຂໍ້ຄວາມເປັນສຽງເວົ້າ (TTS) ເຊັ່ນ YourTTS, ບ່ອນທີ່ເປົ້າຫມາຍແມ່ນເພື່ອສັງເຄາະສຽງເວົ້າຈາກລໍາດັບຂອງ phonemes ແລະ mel-spectrograms ທີ່ສອດຄ້ອງກັນ.

###### vocoder MelGAN ເຊິ່ງເປັນເຄື່ອຂ່າຍປະສາດສຽງ ປະເພດໜຶ່ງທີ່ຖືກອອກແບບມາເພື່ອສ້າງຄື້ນສຽງເວົ້າທີ່ມີຄຸນນະພາບສູງຈາກ mel-spectrograms.

###### Loss ໝາຍເຖິງຄ່າຕ່າງໆທີ່ສູນເສຍ ໃນແບບຈຳລອງເຊັ່ນ: loss\_1 ຫຼື ຄ່າການສູນເສຍຄັ້ງທໍາອິດທີ່ໃຊ້ໃນລະຫວ່າງການຝຶກອົບຮົມຂອງ vocoder MelGAN ເປັນຕົ້ນ

# ທົບທວນເອກະສານ ແລະ ບົດຄົ້ນຄວ້າທີ່ກ່ຽວຂ້ອງ



## ທົບທວນເອກະສານທີ່ກ່ຽວຂ້ອງ

ຜ່ານການຄົ້ນຄວ້າກ່ຽວກັບທິດສະດີ, ວິທີການ, ເອກະສານຄົ້ນຄວ້າທີ່ກ່ຽວຂ້ອງ ແລະ ເຄື່ອງມືຕ່າງໆ ທີ່ນຳໃຊ້ເຂົ້າໃນການຄົ້ນຄວ້າ ການແປຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ ແມ່ນສາມາດສັງລວມບັນດາ ເນື້ອໃນທິດສະດີທີ່ກ່ຽວຂ້ອງດັ່ງລຸ່ມນີ້:



### ການປະມວນຜົນພາສາທຳມະຊາດ (Natural Language Processing)

ການປະມວນຜົນພາສາທຳມະຊາດ (Natural Language Processing) ແມ່ນການວິເຄາະຂໍ້ມູນທີ່ຢູ່ໃນຮູບແບບຂອງຂໍ້ຄວາມຕ່າງໆ ໃຫ້ຄອມພິວເຕີສາມາດເຂົ້າໃຈຂໍ້ຄວາມເຫຼົ່ານັ້ນໄດ້ຄືກັນກັບມະນຸດ (Chowdhury, 2003) ຊຶ່ງປະກອບດ້ວຍຫຼາຍຂັ້ນຕອນຍ່ອຍ ຕັ້ງແຕ່ຂັ້ນຕອນການຈັດກຽມຂໍ້ມູນ ລວມໄປເຖິງການແປງຂໍ້ມູນໃຫ້ຢູ່ໃນຮູບແບບທີ່ຄອມພິວເຕີສາມາດນຳໄປໃຊ້ໃນການປະມວນຜົນຕໍ່ໄດ້ ເຊິ່ງຄອມພິວເຕີສາມາດເຮັດຄວາມເຂົ້າໃຈກັບຂໍ້ມູນທີ່ຢູ່ໃນຮູບແບບຂອງຕົວເລກທີ່ສາມາດນຳໄປຄຳນວນໄດ້ເທົ່ານັ້ນ.

Natural Language Processing ເປັນສ່ວນໜຶ່ງຂອງວຽກງານທາງດ້ານປັນຍາປະດິດ (Artificial Intelligence) ມີຈຸດເລີ່ມຕົ້ນໂດຍເປັນສ່ວນໜຶ່ງຂອງພາສາຄາດ (linguistics) ໃຊ້ໃນການວິເຄາະໃນຮູບແບບຂອງການຕັ້ງກົດ (Rule-Based) (Schank and Abelson, 2013) ຊຶ່ງເປັນການຕັ້ງກົດຂຶ້ນມາໃຊ້ໃນການແກ້ໄຂບັນຫາ ມາປະຍຸກໃຊ້ໃນວຽກງານດ້ານຕ່າງໆ ໂດຍສາມາດສ້າງໄດ້ໂດຍບໍ່ຕ້ອງອາໃສຊຸດຂໍ້ມູນຈຳນວນຫຼາຍ ເຊິ່ງມີຂໍ້ເສຍໃນການວິເຄາະຄຳໃໝ່ໆ ທີ່ເກີດຂຶ້ນມາ ເຮັດໃຫ້ລະບົບບໍ່ຮູ້ຈັກຄຳເຫຼົ່ານັ້ນ ແລະ ຍັງບໍ່ສາມາດຈຳແນກໄດ້ວ່າຈະຕ້ອງມີການເພິ່ມຄຳນັ້ນ ແລະ ກຳນົດປະເພດຂອງຄຳນັ້ນໄວ້ກ່ອນ.

ຈາກການໃຊ້ງານດ້ວຍການຕັ້ງກົດ ຕໍ່ມາໄດ້ມີການປະຍຸກໃຊ້ການຮຽນຮູ້ຂອງເຄື່ອງຈັກ (Machine Learning) ໃນການສ້າງຕົວຈໍາແນກປະເພດຂອງຂໍ້ຄວາມ (Aone et al., 1998). ໂດຍຈະເຮັດວຽກແຕກຕ່າງຈາກການຕັ້ງກົດ ໂດຍວິເຄາະຄຳບາງຄຳ ແລະ ເຮັດການສະຫຼຸບຜົນການວິເຄາະໃນທັນທີເທົ່ານັ້ນ ເຊິ່ງສາມາດສ້າງໄດ້ຈາກການຮວບຮວມຂໍ້ມູນ ແລະ ເຮັດການກຳນົດປະເພດຂໍ້ມູນ ຈາກນັ້ນເຮັດການເຝິກສອນໂດຍໃຊ້ຂັ້ນຕອນວິທີຂອງການຮຽນຮູ້ຂອງເຄື່ອງຈັກ (Machine Learning Algorithm).

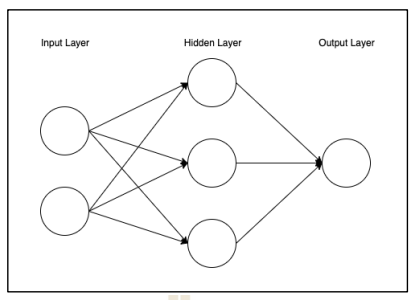
ໃນປັດຈຸບັນການຮຽນຮູ້ຂອງເຄື່ອງຈັກໄດ້ມີການພັດທະນາໄປເປັນການຮຽນຮູ້ຂັ້ນສູງ (Deep Learning) ຊຶ່ງການຮຽນຮູ້ຂັ້ນສູງມີຈຸດເດັ່ນໃນການຮຽນຮູ້ລັກສະນະຕ່າງໆ ໄດ້ຢ່າງຫຼາຍ ແລະ ຮຽນຮູ້ໄດ້ຈຳນວນຫຼາຍ (Goodfellow et al., 2016). ໂດຍສາມາດສ້າງແບບຈໍາລອງໃນການຈຳແນກປະເພດຂອງຂໍ້ຄວາມໄດ້ຢ່າງຊັດເຈນຫຼາຍຍິ່ງຂຶ້ນ ຊຶ່ງໄດ້ມີການນຳມາປະຍຸກໃຊ້ທາງດ້ານ Natural Language Processing ໃນດ້ານຂອງການສ້າງແບບຈໍາລອງສໍາລັບແປພາສາ (Machine Translation) ໂດຍສາມາດແປໄດ້ຢ່າງຖືກຕ້ອງ, ຊັດເຈນ ແລະ ຕໍ່ເນື່ອງ ດ້ວຍຄວາມສາມາດທາງເຕັກໂນໂລຊີ ແລະ ຂັ້ນຕອນວິທີທີ່ສາມາດວິເຄາະຂໍ້ມູນໄດ້ຢ່າງມີປະສິດທິພາບ ເຮັດໃຫ້ມີການພັດທະນາການວິເຄາະທາງດ້ານພາສາທຳມະຊາດເປັນໄປຢ່າງວ່ອງໄວແບບກ້າວກະໂດດ.

### ເຕັກນິກການຮຽນຮູ້ຂັ້ນສູງ (Deep Learning)

ເຕັກນິກການຮຽນຮູ້ຂັ້ນສູງ ເປັນສ່ວນໜຶ່ງຂອງການຮຽນຮູ້ຂອງເຄື່ອງຈັກ (Machine Learning) ຊຶ່ງເປັນຂັ້ນຕອນວິທີ (Algorithm) ທີ່ໃຊ້ສໍາລັບການຮຽນຮູ້ທີ່ສາມາດເຮັດໃຫ້ເຄື່ອງຈັກສາມາດຕັດສິນໃຈໄດ້ຄືກັນກັບມະນຸດ ໂດຍການຮຽນຮູ້ຂອງເຄື່ອງຈັກເປັນການປະຍຸກໃຊ້ຄວາມຮູ້ທາງດ້ານສະຖິຕິ, ໃນການວິເຄາະຂໍ້ມູນ ແລະ ສ້າງແບບຈໍາລອງສຳລັບພະຍາກອນຜົນໄດ້ຮັບຈາກຂໍ້ມູນ.

ຈຸດເລີ່ມຕົ້ນຂອງການຮຽນຮູ້ຂັ້ນສູງ ແມ່ນເລີ່ມມາຈາກໂຄງຂ່າຍປະສາດທຽມ (Neutral Network) ເປັນຂັ້ນຕອນວິທີ (Algorithm) ທີ່ຄິດຄົ້ນຂຶ້ນມາຈາກການຮຽນແບບການເຮັດວຽກຂອງສະໝອງມະນຸດ ຊຶ່ງສະໝອງຂອງມະນຸດມີການເຮັດວຽກທີ່ຊັບຊ້ອນ ແລະ ສາມາດວິເຄາະສິ່ງຕ່າງໆ ເປັນຈຳນວນຫຼາຍໄດ້ຢ່າງມີປະສິດທິພາບ ໂດຍໂຄງຂ່າຍປະສາດທຽມນັ້ນໄດ້ເຮັດການຈຳລອງການເຮັດວຽກຂອງເຊວປະສາດຂຶ້ນມາ ເຊິ່ງແຕ່ລະເຊວກໍມີການເຊື່ອມຕໍ່ເພື່ອສົ່ງຂໍ້ມູນໄປຫາແຕ່ລະເຊວເພື່ອໃຊ້ໃນການຕັດສິນໃຈ ຈຸດເດັ່ນຂອງສະໝອງມະນຸດຄືການທີ່ເຊວປະສາດແຕ່ລະເຊວສາມາດເຊື່ອມຕໍ່ກັນໄດ້ຢ່າງທົ່ວເຖິງ ແລະ ມີການກະຈາຍຕົວໃນການວິເຄາະຂໍ້ມູນໃຫ້ແຕ່ລະເຊວຢ່າງຊັດເຈນ (Hecht-Nielsen, 1992).

ໂຄງຂ່າຍປະສາດທຽມນັ້ນຖືກອອກແບບມາໃຫ້ມີການເຮັດວຽກຄ້າຍກັບສະໝອງມະນຸດ ໂດຍການເຮັດວຽກເບື້ອງຫຼັງຂອງໂຄງຂ່າຍປະສາດທຽມນັ້ນມີໜ່ວຍຍ່ອຍທີ່ເຮັດວຽກຄ້າຍກັບເຊວປະສາດຂອງມະນຸດເອີ້ນວ່າ: Node. ຊຶ່ງ Node ສາມາດລວມຕົວກັນເປັນຈຳນວນໜຶ່ງລຽງຕົວເປັນຊັ້ນ ຈະເອີ້ນວ່າ: Layer ໂດຍແຕ່ລະ Node ຈະມີຂັ້ນຕອນການເຮັດວຽກແບ່ງໜ້າທີ່ຕາມ Layer ເຊັ່ນ: Input Layer, Hidden Layer ແລະ Output Layer ເປັນຕົ້ນ. ທີ່ສະແດງໃນຮູບທີ 2.1 ລຸ່ມນີ້:



ຮູບທີ 2.1 ໂຄງສ້າງໂຄງຂ່າຍປະສາດທຽມໂດຍມີສ່ວນຍ່ອຍໃນການຄໍານວນຄື Node ແລະ ການລຽງຕົວຂອງ Node ເປັນຊັ້ນ ເອີ້ນວ່າ: Layer

ຈາກຮູບທີ 2.1 ການເຮັດວຽກຂອງ Node ມີພື້ນຖານການເຮັດວຽກຈາກທັງໃນສ່ວນຂອງ Linear Regression ໂດຍສາມາດແທນຄ່າໄດ້ດັ່ງ ສົມຜົນທີ (2.1) ໂດຍ Node ນັ້ນຈະມີສ່ວນປະກອບຍ່ອຍເອີ້ນວ່າ: Weight. ຊຶ່ງປຽບທຽບໄດ້ກັບຄ່າ Intercept ຈາກ Linear Regression ທີ່ເອົາໄວ້ໃຊ້ໃນການກໍານົດນໍ້າໜັກຂອງຄ່າຕົວປ່ຽນແຕ່ລະຕົວທີ່ໃຊ້ໃນການວິເຄາະ ແລະ Bias ຊຶ່ງປຽບໄດ້ກັບຄ່າ Coefficient ໃນ Linear Regression ຊຶ່ງປຽບ Node ໄດ້ຄືກັນກັບການເຮັດວຽກຂອງ Linear Regression. ໂດຍໃນຂັ້ນຕອນຂອງການເຝິກສອນນັ້ນ ຈະເຮັດການເຝິກສອນເພື່ອຊອກຫາຄ່າພາລາມິເຕີທີ່ເໝາະສົມກັບຂໍ້ມູນທີ່ໄດ້ເຮັດການນຳມາໃຊ້ເຝິກສອນ ຈາກນັ້ນນຳຄ່າພາລາມິເຕີທີ່ໄດ້ໄປໃຊ້ໃນການສ້າງແບບຈຳລອງເພື່ອໃຊ້ໃນການພະຍາກອນຜົນໄດ້ຮັບ ໂດຍການເຝິກສອນ Neural Network ນັ້ນຈະຕ້ອງເຮັດການເຝິກສອນເປັນຮອບ ໂດຍ 1 ຮອບນັ້ນຈະຕ້ອງເຮັດການເຝິກສອນກັບຂໍ້ມູນທຸກຕົວໃນຂໍ້ມູນທີ່ໃຊ້ສຳລັບເຝິກສອນ ເຊິ່ງເອີ້ນຈຳນວນຮອບໃນການເຝິກສອນວ່າ: Epoch.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 𝑁𝑒𝑢𝑟𝑎𝑙 𝑁𝑒𝑡𝑤𝑜𝑟𝑘(𝑚, 𝑛) = 𝑎𝑐𝑡𝑖𝑣𝑎𝑡𝑖𝑜𝑛(+ ) | | | | (2.1) | |
| ເມື່ອ | | 𝑁𝑒𝑢𝑟𝑎𝑙 𝑁𝑒𝑡𝑤𝑜𝑟𝑘(𝑚, 𝑛) | | ແມ່ນຜົນໄດ້ຮັບຂອງແບບຈຳລອງໂຄງສ້າງໂຄງຂ່າຍປະສາດທຽມ | |
|  | |  | |  | |
|  | | 𝑎𝑐𝑡𝑖𝑣𝑎𝑡𝑖𝑜𝑛 | ແມ່ນຟັງຊັນທີ່ໃຊ້ໃນການແປງຄ່າທີ່ໄດ້ຈາກກາານຄຳນວນຄ່ານ້ຳໜັກຂອງໂຄງສ້າງໂຄງຂ່າຍປະສາດທຽມ | | | | |
|  | |  | ແມ່ນຄ່ານ້ຳໜັກຂອງ Node ໃນໂຄງສ້າງໂຄງຂ່າຍປະສາດທຽມ | | | | |
|  | |  | ແມ່ນຄ່າ Input ຂອງໂຄງສ້າງໂຄງຂ່າຍປະສາດທຽມ | | | | |
|  | |  | ແມ່ນຄ່າ bias ຂອງໂຄງສ້າງໂຄງຂ່າຍປະສາດທຽມ | | | | |

ໂຄງສ້າງໂຄງຂ່າຍປະສາດທຽມນັ້ນມີຢູ່ຫຼາຍຊະນິດ ຊຶ່ງສາມາດແບ່ງອອກໄດ້ດ້ວຍຂັ້ນຕອຍການເຝິກສອນ ໂດຍມີຕົວຢ່າງຂອງປະເພດໂຄງສ້າງໂຄງຂ່າຍປະສາດທຽມທີ່ນິຍົມໃຊ້ເຊັ່ນ: Feedforward Neural Network ທີ່ມີການປ້ອນຂໍ້ມູນຈາກດ້ານໜ້າໄປດ້ານຫຼັງ ໂດຍມີຕົວຢ່າງໄດ້ແກ່ Perceptron ເປັນຕົ້ນ ແລະ Backpropagation Neural Network ທີ່ມີການປ້ອນຂໍ້ມູນ ແລະ ເຮັດການຮຽນຮູ້ໂດຍປ້ອນຄ່າຜົນໄດ້ຮັບທີ່ໄດ້ແບບຍ້ອນກັບ.

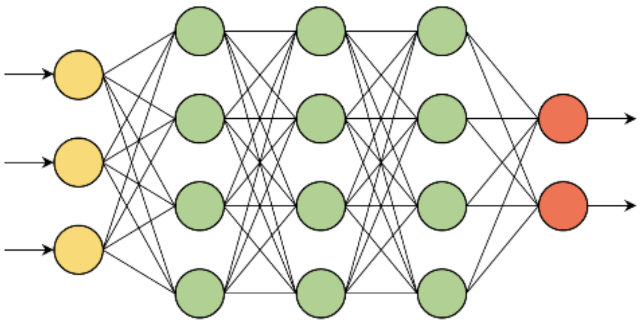
ການຮຽນຮູ້ຂັ້ນສູງເປັນໜຶ່ງໃນຂັ້ນຕອນວິທີ (Algorithm) ທີ່ຖືກພັດທະນາມາຈາກໂຄງຂ່າຍປະສາດທຽມ ຊຶ່ງເປັນການຮຽນແບບການຮຽນຮູ້ຂອງລະບົບປະສາດຂອງມະນຸດ ໂດຍພັດທະນາດ້ວຍການວາງໂຄງສ້າງໃນຮູບແບບວາງຊ້ອນກັນຫຼາຍໆຊັ້ນ ໂດຍມີທັງການວາງຊ້ອນໃນຮູບແບບຂອງປະເພດດຽວກັນໃນທຸກຊັ້ນ ແລະ ການວາງຊ້ອນກັນໂດຍແຕ່ລະຊັ້ນນັ້ນເຮັດວຽກບໍ່ຄືກັນ ເຊິ່ງເກີດຈາກການປະຍຸກໃຊ້ຄວາມສາມາດຂອງແຕ່ລະໂຄງສ້າງທີ່ນຳມາໃຊ້ຮ່ວມກັນເພື່ອເພີ່ມປະສິດທິພາບໃນການວິເຄາະທີ່ດີຍິ່ງຂຶ້ນ (LeCun et al., 2015). ໃນປັດຈຸບັນການຮຽນຮູ້ຂັ້ນສູງ ໄດ້ໃຫ້ປະສິດທິພາບໃນການວິເຄາະ ແລະ ພະຍາກອນຜົນໄດ້ດີກວ່າການໃຊ້ຂັ້ນຕອນວິທີຂອງການຮຽນຮູ້ຂອງເຄື່ອງຈັກ (Machine Learning Algorithm) ແບບດັ່ງເດີມເປັນຢ່າງຫຼາຍ ທັງໃນດ້ານຂອງຈຳນວນຂໍ້ມູນທີ່ເພີ່ມຂຶ້ນຢ່າງວ່ອງໄວ ແລະ ໜ່ວຍປະມວນຜົນຂອງຄອມພິວເຕີທີ່ມີປະສິດທິພາບສູງຂຶ້ນເຮັດໃຫ້ການຮຽນຮູ້ຂັ້ນສູງນັ້ນສາມາດໃຊ້ຕົວປ່ຽນຈຳນວນຫຼາຍໃນການວິເຄາະ ເພື່ອເພີ່ມປະສິດທິພາບໃນການວິເຄາະໄດ້.

ບໍ່ພຽງແຕ່ການເພີ່ມຈຳນວນຊັ້ນເທົ່ານັ້ນ ການຮຽນຮູ້ຂັ້ນສູງໄດ້ທຳການປັບຄວາມສາມາດໃນການຮຽນຮູ້ຂອງຂັ້ນຕອນວິທີ ໂດຍເຮັດການຮຽນຮູ້ແບບ Backpropagation ມາປັບປຸງໂດຍແຍກຂັ້ຍຕອນວິທີທີ່ໃຊ້ໃນການຮຽນຮູ້ອອກເປັນ 2 ຕົວຄື: Loss Function ແລະ Optimize Function ໂດຍ Loss Function ໃຊ້ໃນການຄຳນວນເພື່ອຫາຄ່າຄວາມຜິດພາດທີ່ໄດ້ຈາກການປຽບທຽບລະຫວ່າງຜົນທີ່ໄດ້ຈາກແບບຈຳລອງ ແລະ ຜົນໄດ້ຮັບທີ່ໃຊ້ໃນການເຝິກສອນ ຈາກນັ້ນນຳຄ່າ Loss ທີ່ໄດ້ມາໃຊ້ກັບ Optimize Function ເຊິ່ງເປັນຟັງຊັນສຳລັບການປັບປ່ຽນຄ່າພາລາມິເຕີທີ່ໃຊ້ໃນການຮຽນຮູ້ຂອງແບບຈຳລອງທີ່ສ້າງຂຶ້ນມາ ໂດຍມີການຄິດຄົ້ນ Loss Function ແລະ Optimize Function ຕ່າງໆ ຂຶ້ນມາໃຫ້ເລືອກໃຊ້ໃນປັດຈຸບັນ.

ການຮຽນຮູ້ຂັ້ນສູງນັ້ນໄດ້ມີການຄິດຄົ້ນຂັ້ນຕອນວິທີ ແລະ ໂຄງສ້າງຕ່າງໆ ຂຶ້ນມາ ເພື່ອໃຊ້ໃນການຕອບສະໜອງຕໍ່ຄວາມຕ້ອງການໃນການວິເຄາະຮູບແບບຕ່າງໆ ໂດຍຂັ້ນຕອນວິທີທີ່ໃຊ້ໃນບົດໂຄງການຈົບຊັ້ນນີ້ແມ່ນ Deep Neural Network ແລະ Recurrent Neural Network ເປັນຕົ້ນດັ່ງນີ້:

#### Deep Neural Network (DNN)

ໂຄງຂ່າຍປະສາດຂັ້ນສູງ (Deep Neural Network: DNN) ເປັນການນໍາຂັ້ນຕອນວິທີຂັ້ນສູງໄປໃຊ້ງານໃນຮູບແບບທີ່ງ່າຍທີ່ສຸດ ໂດຍເປັນພັດທະນາມາຈາກໂຄງຂ່າຍປະສາດທຽມດ້ວຍການເພີ່ມຈຳນວນຊັ້ນເຂົ້າໄປຈຳນວນໜຶ່ງ ເຮັດໃຫ້ມີຕົວປ່ຽນທີ່ໃຊ້ສຳລັບການຄັດແຍກຄຸນລັກສະນະທີ່ຫຼາຍຂຶ້ນ ເພື່ອເພີ່ມປະສິດທິພາບໃນການວິເຄາະ ແລະ ພະຍາກອນຜົນໄດ້ຮັບ (Han et al., 2016) ເຊິ່ງໃນການເຮັດວຽກຂອງຕົວໂຄງສ້າງນັ້ນມີລັກສະນະຄືກັນກັບໂຄງຂ່າຍປະສາດທຽມທີ່ສະແດງໃນຮູບທີ 2.2 ລຸ່ມນີ້:



ຮູບທີ 2.2 ຕົວຢ່າງໂຄງສ້າງຂອງຂັ້ນຕອນວິທີ Deep Neural Network

DNN ມີການເຮັດວຽກ ແລະ ໂຄງສ້າງທີ່ຄ້າຍຄືກັນກັບໂຄງຂ່າຍປະສາດທຽມຫຼາຍທີ່ສຸດ ແຕ່ໄດ້ມີການເພີ່ມຈຳນວນຊັ້ນສຳລັບການວິເຄາະຄຸນລັກສະນະຂອງຂໍ້ມູນ ໂດຍຈຳນວນທີ່ຫຼາຍຂຶ້ນເຮັດໃຫ້ສາມາດວິເຄາະຄຸນລັກສະນະໄດ້ລະອຽດຂຶ້ນ ແລະ ເຮັດການເຝິກສອນໂດຍໃຊ້ລັກການຂອງ Backpropagation ໃນການເຝິກສອນ.

DNN ໄດຖືກນຳໄປໃຊ້ໃນການວິເຄາະຂໍ້ມູນໃນຮູບແບບທີ່ມີໂຄງສ້າງຊັດເຈນ ໂດຍໃນການວິເຄາະຂໍ້ມູນຂອງ DNN ນັ້ນ ຈະເຮັດການວິເຄາະໂດຍທີ່ບໍ່ໄດ້ທຳລາຍຕົວຂໍ້ມູນດັ່ງເດີມ ເຮັດໃຫ້ຍັງສາມາດວິເຄາະໄດ້ຢ່າງຊັດເຈນ ແລະ ມີປະສິດທິພາບທີ່ສູງ.

#### Recurrent Neural Network (RNN)

Recurrent Neural Network (RNN) ເປັນຂັ້ນຕອນວິທີຂອງການຮຽນຮູ້ຂັ້ນສູງທີ່ໃຊ້ຫຼັກການວິເຄາະຂໍ້ມູນໃນຮູບແບບຂອງລຳດັບເຫດການ (Sequence) ຊຶ່ງຂໍ້ມູນໃນຮູບແບບນີ້ຈະມີລຳດັບການເກີດຂອງເຫດການທີ່ຊັດເຈນ ແລະ ສາມາດປ່ຽນສະພາບການຂອງເຫດການຕາມລຳດັບໄດ້ ໂດຍຕົວຢ່າງຂໍ້ມູນໃນລັກສະນະນີ້ແມ່ນຂໍ້ມູນຫຸ້ນ ທີ່ມີການເກັບບັນທຶກລາຍວັນ, ຂໍ້ມູນທີ່ຢູ່ໃນຮູບແບບຂອງລຳດັບເວລາ (Time Series) ແລະ ຂໍ້ມູນຂໍ້ຄວາມຕ່າງໆເປັນຕົ້ນ (Mikolov et al., 2010).

RNN ໄດ້ຮັບການພັດທະນາເພີ່ມເຕີມຂຶ້ນກວ່າເກົ່າທີ່ໂຄງຂ່າຍປະສາດທຽມນັ້ນມີການປ້ອນຄ່າ Input ເຂົ້າໄປແລ້ວໄດ້ຄ່າ Output ອອກມາ. ໂດຍ RNN ໄດ້ເຮັດການອອກແບບໃໝ່ໃຫ້ສາມາດນຳຄ່າ Output ໄປຄຳນວນຍ້ອນກັບເປັນ Input ໄດ້ເຊັ່ນກັນ ດັ່ງທີ່ສະແດງໃນຮູບທີ 2.3 ລຸ່ມນີ້:

ຮູບທີ 2.3 ຕົວຢ່າງການເຮັດວຽກໃນເຊວຂອງຂັ້ນຕອນວິທີໂຄງຂ່າຍປະສາດທຽມ RNN

ຈາກຮູບທີ 2.3 ໂຄງຂ່າຍປະສາດທຽມ RNN ນັ້ນຈະມີໂຄງສ້າງທີ່ແຕກຕ່າງຈາກໂຄງຂ່າຍປະສາດທຽມ ໂດຍໂຄງສ້າງຂອງ RNN ທີ່ນ້ອຍທີ່ສຸດຈະເອີ້ນວ່າ: Cell. ໂດຍ Cell ຂອງ RNN ນັ້ນຈະມີທັ່ງໃນສ່ວນຂອງ Hidden State ທີ່ໃຊ້ໃນການຈັດເກັບຂໍ້ມູນການວິເຄາະຈາກຂໍ້ມູນໃນຊຸດທີ່ຜ່ານມາ ເພື່ອໃຊ້ໃນການວິເຄາະໃນຕົວຕໍ່ໄປ.

ໃນການປັບຄ່າພາລາມິເຕີຕ່າງໆນັ້ນ RNN ຈະໃຊ້ວິທີການປ້ອນຂໍ້ມູນແບບຍ້ອນກັບ ເພື່ອເຮັດການປັບຄ່າພາລາມິເຕີຕ່າງໆ ໃນແບບຈຳລອງທີ່ໄດ້ເຮັດການພັດທະນາຂຶ້ນໂດຍໃຊ້ຫຼັກການຂອງກົດລູກໂຊ່ (Chain Rule) ໃນການປັບຄ່າພາລາມິເຕີ. ໂດຍເຮັດການຄຳນວນເພື່ອຫາຄ່າ Gradient ທີ່ໄດ້ຈາກ Error ສຳລັບການປັບຄ່າພາລາມິເຕີຂອງແບບຈຳລອງ ດັ່ງສົມຜົນທີ 2.2 ແລະ 2.3 (Premjith et al., 2018).

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | | (2.2) |
|  |  | | (2.3) |
| ເມື່ອ |  | ແມ່ນາ Hidden State ຂອງໜ່ວຍເວລາ | |
|  |  | ແມ່ນຄ່າຟັງຊັນສຳລັບຄຳນວນຄ່າ Hidden State ຂອງໜ່ວຍເວລາ | |
|  |  | ແມ່ນຄ່ານ້ຳໜັກສຳລັບຄຳນວນຄ່າ Input ໃນ Hidden State | |
|  |  | ແມ່ນຄ່າ Input ຂອງຂໍ້ມູນ | |
|  |  | ແມ່ນຄ່ານ້ຳໜັກສຳລັບຄຳນວນຄ່າ Hidden ຈາກໜ່ວຍເວລາກ່ອນໜ້າໃນ Hidden State | |
|  |  | ແມ່ນຄ່າ Hidden State ຈາກໜ່ວຍເວລາກ່ອນໜ້າ | |
|  |  | ແມ່ນຄ່າ Bias ສຳລັບຄຳນວນຕ່າ Hidden State | |
|  |  | ແມ່ນຄ່າ Output ທີ່ໄດ້ຈາກແບບຈຳລອງ RNN | |
|  |  | ແມ່ນຟັງຊັນສຳລັບຄຳນວນຄ່າ Output | |
|  |  | ແມ່ນຄ່ານ້ຳໜັກສຳລັບຄຳນວນຄ່າ Hidden State | |
|  |  | ແມ່ນຄ່າ Bias ສຳລັບຄຳນວນຄ່າ Output | |

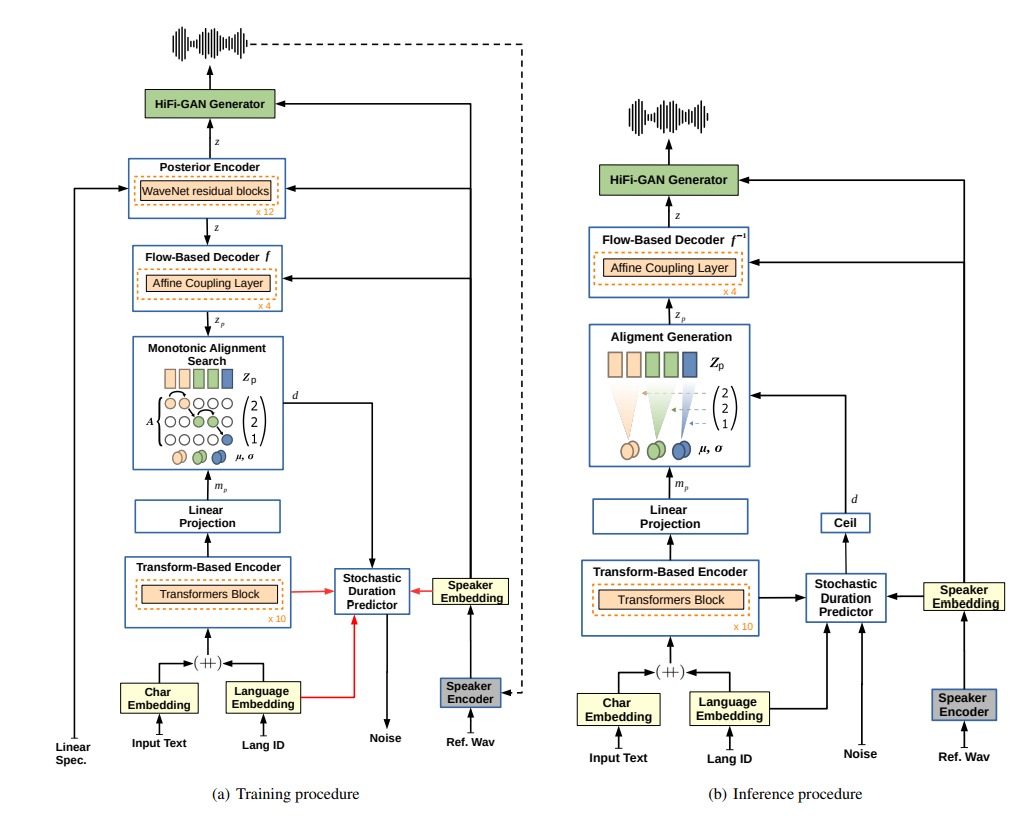
ຈາກສົມຜົນທີ (2.2) ໃນການປັບຄ່າພາລາມິເຕີໃນສ່ວນຂອງ Hidden State ນັ້ນໄດ້ເກີດປັນຫາຂຶ້ນຈາກການໃຊ້ງານໂຄງຂ່າຍປະສາດແບບ RNN ໄດ້. ເນື້ອງຈາກ, ການກຳນົດລຳດັບຂອງຂໍ້ມູນທີ່ມີຂະໜາດທີ່ຍາວເກີນໄປ ເຮັດໃຫ້ຄ່າ Gradient ມີຄ່າຕ່ຳລົງຈົນບໍ່ສາມາດເຫັນການປ່ຽນແປງໄດ້ ຈຶ່ງເຮັດໃຫ້ໂຄງຂ່າຍປະສາດແບບ RNN ນັ້ນມີຂໍ້ຈໍາກັດໃນການໃຊ້ງານວິເຄາະຂໍ້ມູນທີ່ມີຄວາມຍາວທີ່ຫຼາຍເກີນໄປ.

ໂຄງຂ່າຍປະສາດແບບ RNN ໄດ້ຮັບຄວາມນິຍົມໄປໃຊ້ງານກັບຂໍ້ມູນທີ່ໃນຮູບແບບຂອງຂໍ້ຄວາມເປັນຈຳນວນຫຼາຍ ໂດຍສະເພາະການປະຍຸກໃຊ້ທາງດ້ານຂອງ Language Modeling ແລະ Machine Translation ເຊິ່ງເປັນເພາະການເບິ່ງຂໍ້ຄວາມໃຫ້ຢູ່ໃນຮູບແບບຂໍ້ມູນແບບລຳດັບ ແລະ ເຮັດການວິເຄາະໂດຍຄຳນຶງເຖິງຕຳແໜ່ງຂອງ Input ເຮັດໃຫ້ RNN ສາມາດວິເຄາະຂໍ້ມູນຂໍ້ຄວາມໄດ້ຢ່າງມີປະສິດທິພາບ.

### ເຕັກນິກຂອງ YourTTS

YourTTS ສ້າງຂຶ້ນຕາມ VITS ແຕ່ປະກອບມີ novel modification algorithm ສຳລັບການເຝິກແບບຈຳລອງທີ່ເປັນແບບ Zero-Shot Multilingual ແລະ ແບບຫຼາຍພາສາໃນສຽງດຽວ. ເຊິ່ງຈະບໍ່ຄືກັບ VITS ທີ່ຈະໃຊ້ Phonemes, ໃນຮູບແບບຂອງພວກເຮົາ, ພວກເຮົາໃຊ້ຂໍ້ຄວາມດິບເປັນການປ້ອນຂໍ້ມູນແທນ phonemes. ໂດຍທີ່ຜົນໄດ້ຮັບທີ່ໄດ້ຈະຍິ່ງຊັດເຈນຄືກັບຕົວຈິງຫຼາຍຂຶ້ນເມື່ອໃຊ້ກັບພາສາທີ່ບໍ່ຄ່ອຍຈະເປັນ Open-Source ເຊິ່ງຈະແປງ graphem-to-phoneme.

YourTTS ໃນລະຫວ່າງການເຝິກສອນ ແລະ ການອະນຸມານ ແມ່ນສະແດງໃຫ້ເຫັນໃນຮູບທີ່ 2.4, ບ່ອນທີ່ (++) ເປັນຕົວຊີ້ໃຫ້ເຫັນເຖິງ ການເຊື່ອມຕໍ່ກັນຂອງຂໍ້ມູນ, ການເຊື່ອມຕໍ່ສີແດງໝາຍຄວາມວ່າ: ການເຊື່ອມຕໍ່ນີ້ຈະແຜ່ຂະຫຍາຍແບບບໍ່ໄລລະດັບສີ ແລະ ເສັ້ນຂີດຍະຫວ່າງ (--) ການເຊື່ອມຕໍ່ເປັນທາງເລືອກ. ເຊິ່ງພວກເຂົາບໍ່ຈໍາແນກເຄືອຂ່າຍ Hifi-GAN ເພື່ອຄວາມງ່າຍດາຍ ແລະ ເພື່ອເຮັດໃຫ້ຄວາມສາມາດໃນການສ້າງແບບຈຳລອງແບບ Zero Multi-Speaker, ພວກເຂົາກຳນົດ Layer ການເຊື່ອມຕໍ່ທີ່ສຳພັນກັນທັງໝົດດ້ວຍ flowbased decoder, the posterior encoder ແລະ ຕົວຖອດລະຫັດສຽງທີ່ຢູ່ໃນລຳໂພງ, ພວກເຂົາໃຊ້ການປັບສະພາບໂດຍລວມ ໃນບລ໋ອກທີ່ເຫຼືອຂອງ Layer ຂໍ້ຕໍ່ ແລະ ຕົວເຂົ້າລະຫັດແບບ posterior. ພວກເຂົາຍັງສະຫຼຸບເຖິງຕົວຖອດລະຫັດສຽງທີ່ຢູ່ໃນລຳໂພງ ກັບ ສິ່ງທີ່ຂໍ້ຄວາມທີ່ເຂົ້າລະຫັດສົ່ງອອກມາ ແລະ ກັບ ສິ່ງທີ່ຂໍ້ຄວາມທີ່ເຂົ້າລະຫັດສົ່ງອອກມາຂອງກ່ອນໜ້ານີ້. ພວກເຂົາສົ່ງພວກມັນໄປຫາຕົວຄາດຄະເນໄລຍະເວລາ ແລະ ຕົວສົ່ງສຽງ ຕາມລໍາດັບ. ພວກເຂົາໃຊ້ Layer ການຄາດຄະເນແບບເສັ້ນຊື່ເພື່ອໃຫ້ກົງກັບຂະໜາດກ່ອນການສະຫຼຸບອົງປະກອບ (Casanova, Weber, Shulby, Junior, Golge and Antonelli, 2023) ດັ່ງນີ້:



ຮູບທີ 2.4 ແຜນວາດການເຝິກສອນແບບຈໍາລອງ ແລະ ອະນຸມານແບບຈຈໍາລອງ YourTTS

### Google Speech-To-Text API

Google Speech-To-Text API ປະກອບມີ 3 ວິທີຫຼັກໃນການສ້າງແບບຈຳລອງການຈື່ຈຳສຽງເວົ້າ (Google Speech-To-Text API, 2023) ຄື:

Synchronous Recognition (REST and gRPC)

###### Asynchronous Recognition (REST and gRPC)

###### Streaming Recognition (gRPC only)

ສຳລັບບົດໂຄງການຈົບຊັ້ນສະບັບນີ້ ພວກຂ້າພະເຈົ້າໄດ້ນຳໃຊ້ວິທີການ Streaming Recognition (gRPC only) ໃນການສ້າງລະບົບການແປສຽງເວົ້າໄປຂໍ້ຄວາມ ມີວິທີການດັ່ງນີ້:

#### ການຮ້ອງຂໍ API ການຈື່ຈຳສຽງໄປເປັນຂໍ້ຄວາມແບບ Streaming (Streaming Speech-to-Text API Recognition Requests)

ການເອີ້ນໃຊ້ API ການແປຄຳເວົ້າໄປເປັນຂໍ້ຄວາມແບບ Streaming ໄດ້ຮັບການອອກແບບມາສຳລັບການຈັບຮູບພາບ (ຫຼື ຄື່ນສຽງ) ແລະ ການຮຽນຮູ້ສຽງແບບ Real-Time ພາຍໃນການ Streaming ແບບສອງທິດທາງ. ແອບພິເຄເຊີນສາມາດສົ່ງຄໍາຮ້ອງຂໍເປັນສຽງເພື່ອສົ່ງແບບ Streaming ແລະ ຮັບຜົນທີ່ໄດ້ຮັບແບບຊົ່ວຄາວ ແລະ ສຸດທ້າຍເມື່ອຖືກຕອບກັບມາແບບ Real-Time Streaming ຜົນໄດ້ຮັບລະຫວ່າງການສະແດງຜົນຈົນໄປຮອດຜົນໄດ້ຮັບການຈື່ຈໍາສ່ວນຊ່ອງສຽງທີ່ຖືກສົ່ງອອກໄປລ່າສຸດ ໃນຂະນະທີ່ຜົນໄດ້ຮັບການຈື່ຈຳຂັ້ນສຸດທ້າຍສະແດງອອກມາຈົນໄປຮອດການພະຍາກອນສຽງສຸດທ້າຍທີ່ດີທີ່ສຸດອອກມາ.

#### ການຮ້ອງຂໍແບບ Streaming

ການເອີ້ນໃຊ້ Speech API ແບບ Streaming ຈຳເປັນຕ້ອງສົ່ງຄຳຂໍຫຼາຍໆລາຍການເປັນຕົ້ນແມ່ນ StreamingRecognizeRequest ຕ້ອງມີການກຳນົດຄ່າປະເພດໃຫ້ກັບ StreamingRecognizeRequest ໂດຍບໍ່ມີສຽງປະກອບ. ຫຼັງຈາກນັ້ນ, ແມ່ນນຳເອົາ StreamingRecognizeRequest ທີ່ສົ່ງຜ່ານ Streaming ດຽວກັນມາປະກອບເປັນ Frame ທີ່ຕໍ່ເນື່ອງກັນຂອງສຽງດິບທີ່ເປັນ Bytes.

StreamingRecognitionConfig ປະກອບດ້ວຍ fields ດັ່ງນີ້:

###### Config- (ຈໍາເປັນ) ຈະມີການກໍານົດຄ່າສໍາລັບສຽງປະເພດ RecognitionConfig ແລະ ຄືກັນກັບທີ່ສະແດງພາຍໃນຄໍາຂໍຂອງ Synchronous ແລະ Asynchronous

###### single\_utterance- (ບໍ່ບັງຄັບ ແຕ່ຄ່າເລີ່ມຕົ້ນເປັນ false) ຈະລະບຸວ່າຄຳຂໍນີ້ຄວນສິນສຸດລົງໂດຍອັດຕະໂນມັດ ຫຼື ບໍ່ ຫຼັງຈາກທີ່ກວດບໍ່ເຫັນຄຳເວົ້ານີ້ອີກຕໍ່ໄປ ແຕ່ວ່າຖ້າຕັ້ງຄ່າໄວ້ການແປງຄຳເວົ້າໄປເປັນຂໍ້ຄວາມຈະກວດຫາການຢຸດເປັນຈັງຫວະ, ຄວາມງຽບ ຫຼື ສຽງທີ່ບໍ່ແມ່ນຄຳເວົ້າ ເພື່ອເປັນການກຳນົດວ່າເມື່ອໃດຈະສິ້ນສຸດການຈື່ຈຳ ແຕ່ຖ້າບໍ່ໄດ້ຕັ້ງຄ່າໄວ້ແບບຈຳລອງຈະຟັງ ແລະ ປະມວນຜົນຕໍ່ໄປຈົນກວ່າຈະປິດແບບຈຳລອງ ຫຼື ເກີນຂີດຈຳກັດການເອີ້ນໃຊ້ແບບຈຳລອງ. ການຕັ້ງຄ່າໃຫ້ single\_utterance ໃຫ້ເປັນຄ່າ true ຈະເປັນປະໂຫຍດສໍາລັບການປະມວນຜົນຄຳສັ່ງສຽງ.

###### interim\_results- (ບໍ່ບັງຄັບ ຄ່າເລີ່ມຕົ້ນເປັນ false) ຈະລະບຸວ່າຄໍາຂໍ Streaming ຄວນສົ່ງຄືນຜົນໄດ້ຮັບຊົ່ວຄາວທີ່ອາດຈະຖືກປັບປຸງໃນພາຍຫຼັງ (ຫຼັງຈາກປະມວນຜົນສຽງເພີ່ມເຕີມແລ້ວ) ຖ້າເກີດວ່າຕັ້ງຄ່າໃຫ້ is-final ເປັນ false ຈະເຮັດໃຫ້ຜົນໄດ້ຮັບທີ່ປະມວນອອກມາລ່າສຸດຈະຖືກບັນທຶກໄວ້.

### Unified Modeling Language (UML)

UML (Unified Modeling Language) ເປັນພາສາຮູບພາບ ຫຼື ສັນຍາລັກທີ່ໃຊ້ເພື່ອສະແດງຄວາມຄິດຂອງເຮົາທີ່ມີຕໍ່ລະບົບອອກມາເປັນແຜນພາບ, ເຊິ່ງປະກອບໄປດ້ວຍຮູບພາບ ຫຼືສັນຍາລັກ, ເພື່ອໃຊ້ໃນການວິເຄາະ ແລະ ອອກແບບລະບົບແບບວັດຖຸ ໂດຍສະເພາະໃນລະບົບທີ່ມີຂະໜາດໃຫຍ່ ແລະ ມີຄວາມຊັບຊ້ອນ ໃຫ້ສາມາດເຂົ້າໃຈໄດ້ງ່າຍຂຶ້ນ.

UML Structure

Things

1. Structure Things:

* Class
* Interface
* Collaboration
* Use Case
* Activity Class
* Component
* Node

1. Behavioral Things:

* Interaction
* State Machine

1. Grouping Things:

* Package

1. Annotation Things:

* Note

Diagram

1. Structure Diagram:

* Class Diagram
* Object Diagram
* Package Diagram
* Composite Diagram
* Component Diagram
* Deployment Diagram

1. Behavioral Diagram:

* Use Case Diagram
* Sequence Diagram
* Interaction Overview Diagram
* State Chart Diagram
* Activity Diagram
* Communication Diagram
* Timing Diagram

Relationships

ຮູບທີ 2.5 ຮູບພາບສະແດງອົງປະກອບຂອງ UML

#### Use Case Diagram

ການສ້າງ Use Case Diagram ແມ່ນການແນໃສ່ການສ້າງ Functional Requirement ເຊິ່ງໝາຍເຖິງການວິເຄາະຄວາມຕ້ອງການຂອງລະບົບດ້ານຄວາມສາມາດ ຫຼື ໜ້າທີ່ ທີ່ລະບົບຈະຕ້ອງເຮັດໄດ້. ຕໍ່ຈາກນັ້ນ, ຈຶ່ງນໍາເອົາ Use Case Diagram ໄປເປັນພື້ນຖານໃນການສ້າງແຜ່ນພາບຊະນິດອື່ນໆທີ່ເປັນຄວາມສົນໃຈດ້ານ Non-Functional Requirement ເພີ່ມເຕີມໃນຂັ້ນຕໍ່ໄປ ຈົນໃນທີ່ສຸດກໍສາມາດສະແດງໃຫ້ເຫັນເຖິງໂຄງສ້າງຂອງລະບົບໃໝ່ໄດ້ຄົບທຸກສ່ວນ (ພານິດາ ພານິດຊະກຸນ, 2009).

##### ກ. Use Case: ແມ່ນໜ້າທີ່ ທີ່ລະບົບຈະຕ້ອງໄດ້ເຮັດ (Functionality) ໂດຍລະບົບດັ່ງກ່າວຕ້ອງເປັນລະບົບທີ່ເຮົາສົນໃຈ. ໃຊ້ສັນຍາລັກຮູບແອນລິບ ພ້ອມທັງຂຽນຊື່ Use Case ໄວ້ດ້ານລຸ່ມ, ໂດຍໃຊ້ຄໍາກຳມະ ຫຼື noun phase ກໍໄດ້.

ຮູບທີ 2.6 ຮູບພາບສະແດງສັນຍາລັກ Use Case

##### ຂ. Actor: ແມ່ນຜູ້ທີ່ກ່ຽວຂ້ອງກັບລະບົບລວມທັງ Primary Actor ແລະ Stakeholder Actor ທີ່ສັນຍາລັກຮູບຄົນ ພ້ອມທັງຂຽນຊື່ໄວ້ດ້ານລຸ່ມ, ແຕ່ຖ້າ Actor ຫາກເປັນລະບົບອື່ນທີ່ຢູ່ນອກລະບົບທີ່ເຮົາສົນໃຈ ຈະໃຊ້ຮູບສີ່ຫຼ່ຽມແລ້ວຂຽນຄໍາວ່າ Actor ດັ່ງຮູບ:

****

ຮູບທີ 2.7 ຮູບພາບສະແດງສັນຍາລັກ Actor

##### ຄ. System Boundary: ແມ່ນເສັ້ນແບ່ງລະຫວ່າງລະບົບກັບຜູ້ກະທຳຕໍ່ລະບົບ (Use Case ກັບ Actor) ໃຊ້ຮູບສີຫຼ່ຽມເປັນສັນຍາລັກ ພ້ອມທັງຂຽນຊື່ລະບົບໄວ້ດ້ານເທິງ.

System Name

ຮູບທີ 2.8 ຮູບພາບສະແດງສັນຍາລັກ System Boundary

##### ງ. Relationship: ແມ່ນຄວາມສໍາພັນລະຫວ່າງ Use Case, ສັນຍາລັກດ້ວຍ <<…>>

###### Include Relationship: ເປັນຄວາມສໍາພັນລະຫວ່າງ Use Case ລັກສະນະເອີ້ນໃຊ້ ເກີດຂຶ້ນເມື່ອ Use Case ໜຶ່ງໄປເອີ້ນໃຊ້ Use Case ອື່ນ ເພື່ອເຮັດການປະມວນຜົນໃນ Use Case ຕົນເອງສໍາເລັດ.

<<include>>

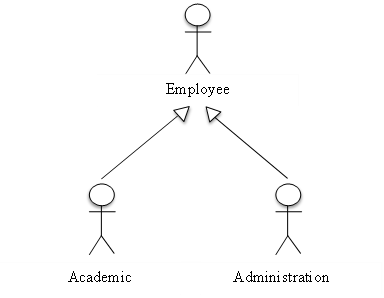
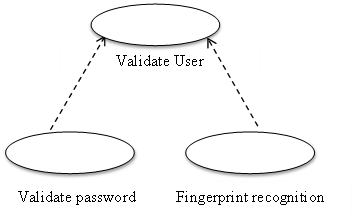
ຮູບທີ 2.9 ຮູບພາບສະແດງຄວາມສໍາພັນ Include Relationship

###### Extend Relationship: ເປັນຄວາມສໍາພັນລະຫວ່າງ Use Case ແບບເພີ່ມເຕີມ ຫຼື ໜ້າທີ່ພິເສດເພີ່ມ (ບໍ່ໄດ້ເກີດຂຶ້ນເປັນປະຈໍາ).

* << extend >>

ຮູບທີ 2.10 ຮູບພາບສະແດງຄວາມສໍາພັນ Extend Relationship

##### ຈ. Generalization / Specialization Relationship: ແມ່ນແນວຄິດທີ່ໃຊ້ສະແດງຄວາມສໍາພັນແບບຈໍາແນກປະເພດຂອງ Use Case ຫຼື Actor ໂດຍ Use Case / Actor ທີ່ເປັນຫຼັກຂອງການຈໍາແນກປະເພດ ຈະເອີ້ນວ່າ “Parent” ສ່ວນ Use Case / Actor ທີ່ຖືກຈຳແນກ ຈະເອີ້ນວ່າ “Child” ສໍາລັບສັນຍາລັກທີ່ໃຊ້ຈະເປັນເສັ້ນຊື່ມີຫົວລູກສອນລາກຈາກ Child ຊີ້ໄປທີ່ Parent.



ຮູບທີ 2.11 ຮູບພາບສະແດງຄວາມສໍາພັນ Generalization

#### Sequence Diagram

Sequence Diagram ເປັນແຜ່ນພາບທີ່ໃຊ້ສະແດງໃຫ້ເຫັນວ່າ Object ໃນລະບົບມີການໂຕ້ຕອບກັນແບບໃດ, ໂດຍຈະເນັ້ນການສົ່ງ Message ລະຫວ່າງ Object. ສັນຍາລັກທີ່ໃຊ້ໃນ Sequence Diagram ປະກອບມີ:

###### Actor: ຄືຜູ້ກ່ຽວຂ້ອງກັບລະບົບ ຫຼື ຜູ້ກະທຳຕໍ່ລະບົບ. ສັນຍາລັກດ້ວຍຮູບຄົນ(Stick Man Icon)

Actor

ຮູບທີ 2.12 ຮູບພາບສະແດງສັນຍາລັກຂອງ Actor ໃນ Sequence Diagram

###### Object: ຄືວັດຖຸທີ່ຕ້ອງເຮັດໜ້າທີ່ຕອບສະໜອງ. ສັນຍາລັກດ້ວຍຮູບສີ່ແຈສາກພ້ອມທັງຂຽນຊື່ Object ກຳກັບໄວ້ດ້ານໃນ

objectName:className

Object

ຮູບທີ 2.13 ຮູບພາບສະແດງສັນຍາລັກຂອງ Object ໃນ Sequence Diagram

###### Lifeline: ແມ່ນເສັ້ນຊ່ວງຊີວິດ/ຊ່ວງເວລາເຮັດວຽກຂອງ Object ຫຼື Class. ສັນຍາລັກດ້ວຍ:

Lifeline

ຮູບທີ 2.14 ຮູບພາບສະແດງສັນຍາລັກຂອງ Lifeline ໃນ Sequence Diagram

###### Focus of control/Activation: ແມ່ນຈຸດເລີ່ມຕົ້ນ ແລະ ສິ້ນສຸດຂອງແຕ່ລະກິດຈະກຳໃນລະຫວ່າງທີ່ວັດຖຸເຮັດວຽກຢູ່. ສັນຍາລັກດ້ວຍ:

Focus of control

ຮູບທີ 2.15 ຮູບພາບສະແດງສັນຍາລັກຂອງ Activation ໃນ Sequence Diagram

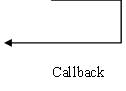
###### Message:ແມ່ນຄຳສັ່ງ ຫຼື Function ທີ່ Class ໜຶ່ງສົ່ງໃຫ້ອີກ Class ໜຶ່ງ ແລະ ສາມາດສົ່ງກັບໄດ້. ສັນຍາລັກດ້ວຍ:

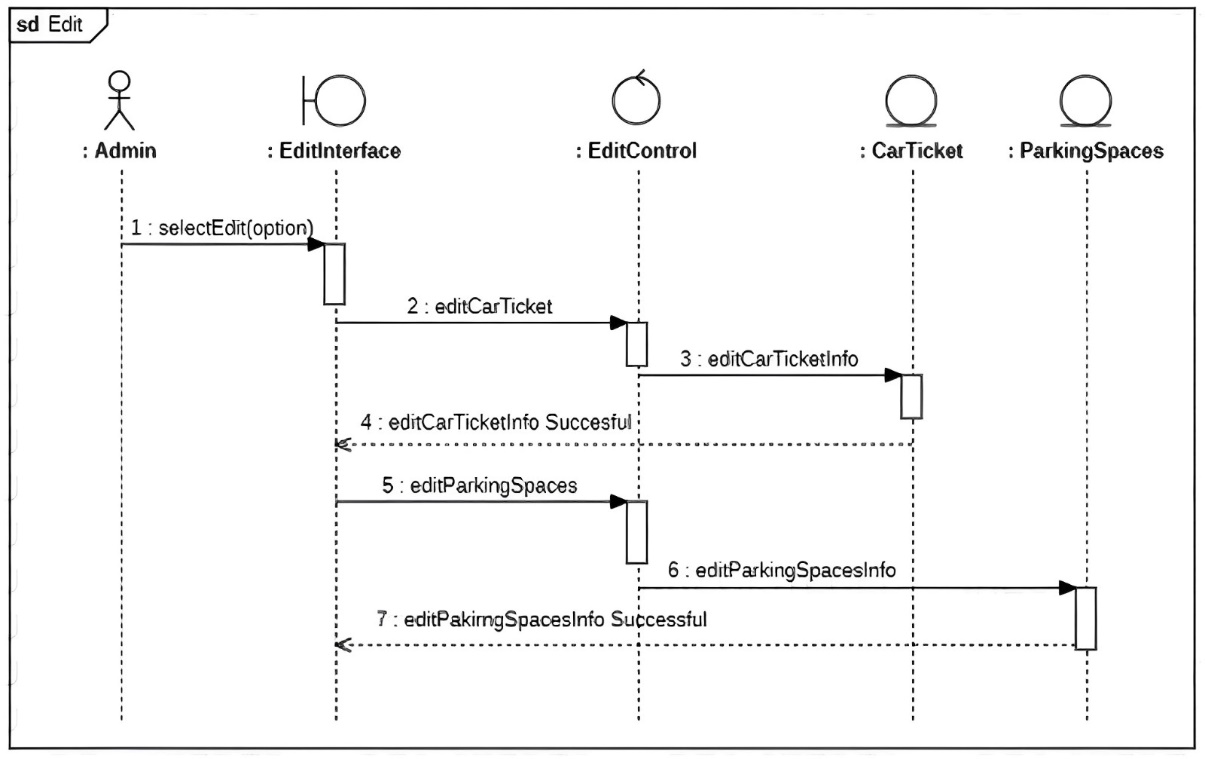
Message

Message()

ຮູບທີ 2.16 ຮູບພາບສະແດງສັນຍາລັກຂອງ Message ໃນ Sequence Diagram

###### Callback/Self Delegation: ແມ່ນການປະມວນຜົນ ແລະ ຄືນຄ່າທີ່ໄດ້ພາຍໃນ Object ດຽວກັນກັບຄືນ. ສັນຍາລັກດ້ວຍ:

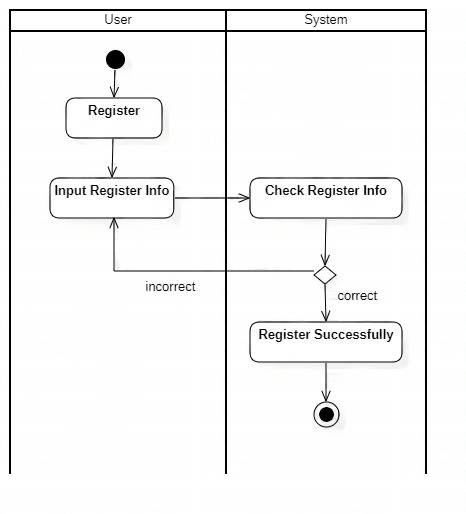
****

ຮູບທີ 2.17 ຮູບພາບສະແດງສັນຍາລັກຂອງ Callback ໃນ Sequence Diagram

ຮູບທີ 2.18 ຮູບພາບສະແດງ Sequence Diagram

#### Activity Diagram

ການສ້າງ Activity Diagram ແມ່ນເລີ່ມຈາກການເລືອກກິດຈະກຳ ຫຼື Method ທີ່ຕ້ອງການອະທິບາຍ ຫຼື ສ້າງ Activity ສໍາລັບທຸກໆ Use Case, ຫຼັງຈາກນັ້ນກໍຂຽນວາດລໍາດັບກິດຈະກຳ ແລະ ແຕ້ມແຜນວາດລໍາດັບກິດຈະກຳ.



ຮູບທີ 2.19 ຮູບພາບສະແດງ Activity Diagram

### ຄວາມຮູ້ກ່ຽວກັບ Flowchart

Flowchart ແມ່ນແຜນວາດທີ່ສະແດງໃຫ້ເຫັນເຖິງແນວຄິດ ແລະ ຂັ້ນຕອນການເຮັດວຽກຂອງໂປຣແກຣມ ເຊິ່ງຊ່ວຍສະແດງໃຫ້ເຫັນເຖິງພາບລວມຂອງໂປຣແກຣມ ເຮັດໃຫ້ສາມາດຂຽນໂປຣແກຣມໄດ້ງ່າຍຂຶ້ນ. ການຂຽນ Flowchart ນັ້ນຈະຕ້ອງໃຊ້ສັນຍາລັກຕ່າງໆແທນຄຳອະທິບາຍ ແລະ ຂະບວນການເຮັດວຽກຂອງໂປຣແກຣມໃນແຕ່ລະສ່ວນຕັ້ງແຕ່ເລີ່ມຕົ້ນຈົນເຖິງສິ້ນສຸດການເຮັດວຽກຂອງໂປຣແກຣມ ເພື່ອໃຫ້ຜູ້ພັດທະນາໂປຣແກຣມໄດ້ເຂົ້າໃຈແນວຄິດ ແລະ ການເຮັດວຽກຂອງທີ່ຊັດເຈນຫຼາຍທີ່ສຸດ (IT-HR, 2014).

ຕາຕະລາງທີ 2.1 ຕາຕະລາງສະແດງສັນຍາລັກທີ່ໃຊ້ແຕ້ມແຜນວາດ Flowchart

|  |  |
| --- | --- |
| **ສັນຍາລັກ** | **ຄວາມໝາຍ** |
|  | ຈຸດເລີ່ມຕົ້ນ ຫຼື ຈຸດສິ້ນສຸດ |
|  | ຮັບຂໍ້ມູນ ຫຼື ສະແດງຂໍ້ມູນໂດຍບໍ່ລະບຸຊະນິດອຸປະກອນ |
|  | ຂະບວນການ ແລະ ການຄຳນວນ |
|  | ການຕັດສິນໃຈ ຫຼື ການປຽບທຽບ |
|  | ຈຸດເຊື່ອມຕໍ່ |
|  | ລູກສອນສະແດງທິດທາງຄວາມສຳພັນຂອງການເຮັດວຽກໃນລະບົບວຽກງານ ຫຼື ລຳດັບວຽກງານ |

## ບົດຄົ້ນຄວ້າທີ່ກ່ຽວຂ້ອງ

ການຮຽນຮູ້, ການອ່ານ ແລະ ການຈື່ຈຳ ຂໍ້ຄວາມພາສາລາວຈາກເອກະສານຮູບພາບ ດ້ວຍເຕັກນິກຂອງປັນຍາປະດິດຂັ້ນສູງ ເປັນແບບຈຳລອງທີ່ຖືກພັດທະນາໂດຍ ທ້າວ ຄຳສະໄໝ ໄຊຍະວົງ (2022) ລາວໄດ້ສຶກສາ Lao Optical Character Recognition (LOCR) ເປັນການຈຳແນກ ແລະ ຈື່ຈຳພາສາລາວ ຈາກຮູບພາບ ໂດຍໃຊ້ວິທີການ CNN (Convolution Neural Network) ເຊິ່ງແມ່ນ Algorithm ທີ່ຈັດຢູ່ໃນການຮຽນຮູ້ຂັ້ນສູງ (Deep learning) ໂດຍຂັ້ນຕອນການດໍາເນີນການສ້າງແບບຈໍາລອງນັ້ນ ປະກອບໄປດ້ວຍການກະກຽມຂໍ້ມູນຮູບພາບຕົວອັກສອນ (Dataset), ການກະກຽມກ່ອນການປະມວນຜົນ ແລະ ຂັ້ນຕອນການປະມວນຜົນ. ສຳລັບການກະກຽມຂໍ້ມູນແມ່ນໄດ້ເກັບຈຳນວນຕົວເລກສາກົນ 10 ຕົວ ແລະ ຕົວອັກສອນລາວ 53 ຕົວ ລວມເປັນ 63 ຕົວ ເຊິ່ງແຕ່ລະຕົວມີຢູ່ 1,680 ຮູບພາບ ເຊິ່ງລວມຂໍ້ມູນຮູບພາບຕົວເລກ ແລະ ຕົວອັກສອນລາວທັງໝົດມີ 105,840 ຮູບພາບ. ຂັ້ນຕອນການປະມວນຜົນ ແມ່ນໄດ້ນໍາໃຊ້ວິທີການຂອງ CNN ເພື່ອຮຽນຮູ້ຕົວອັກສອນຈາກຮູບພາບ ເຊິ່ງໄດ້ນຳເອົາໂຄງສ້າງສະຖາປັດຕະຍະກຳ (Architecture) ຂອງ CNN ນຳມາປະມວນຜົນມີ LeNet-5, AlexNet ແລະ Improved VGG16 ທີ່ໄດ້ປັບປຸງໂຄງສ້າງແລ້ວ. ໂດຍໄດ້ນໍາເອົາຂໍ້ມູນເພື່ອອົບຮົມ (Train) ຈຳນວນ 67,737 ຮູບພາບ, ຈຳນວນທີ່ນຳໄປກວດສອບຄວາມຖືກຕ້ອງ (Validation) ມີ 16,935 ຮູບພາບ (ຂໍ້ມູນອົບຮົມ ແລະ ກວດສອບຄວາມຖືກຕ້ອງ 80%). ຈໍານວນທີ່ນໍາມາທົດສອບ (Test) ມີ 21,168 ຮູບພາບ (ຂໍ້ມູນທົດສອບ 20%), ເຊິ່ງຜົນຂອງການຮຽນຮູ້ຂອງແຕ່ລະໂຄງສ້າງສະຖາປັດຕະກໍາພົບວ່າ: LeNet-5 ແມ່ນມີຄວາມຖືກຕ້ອງ 90,07%, AlexNet ແມ່ນມີຄວາມຖືກຕ້ອງ 94,52% ແລະ Improved VGG16 ແມ່ນມີຄວາມຖືກຕ້ອງສູງເຖິງ 97,80%.

ການຈື່ຈຳຕົວອັກສອນໄຕທຳ ແລະ ພັດທະນາລະບົບດ້ວຍເຄືອຂ່າຍປະສານທຽມແບບຄອນໂວລູເຊີນ ເປັນແບບຈຳລອງທີ່ຖືກພັດທະນາໂດຍ ທ້າວ ຍີທອນ ອິ່ນແສງໃສ (2022) ລາວໄດ້ສຶກສາ ການຈື່ຈຳຕົວອັກສອນດ້ວຍແສງ (OCR) ເປັນວິທີການປ່ຽນຕົວອັກສອນທີ່ຢູ່ໃນຮູບພາບໃຫ້ເປັນຕົວອັກສອນທີ່ສາມາດແກ້ໄຂໄດ້ໃນຄອມພິວເຕີທີ່ນິຍົມໃຊ້ໃນປະຈຸບັນ. ເປັນການສຶກສາຕົວອັກສອນໄຕທໍາ, ເຕັກນິກການຈື່ຈຳຕົວອັກສອນ ແລະ ການສ້າງແບບຈຳລອງ ຫຼື ແມ່ແບບ ANN, CNN ແລະ KNN ໂດຍການນຳໃຊ້ວິທີການຂອງການປະມວນຜົນຮູບພາບ, ການຮຽນຮູ້ດ້ວຍຄອມພິວເຕີ (Machine Learning), Deep Leaming ແລະ Computer Vision. ໃນການຝຶກສອນເຄືອຂ່າຍປະສາດທຽມໃຫ້ຈື່ຈຳຕົວອັກສອນໄຕ ແມ່ນປະກອບດ້ວຍຮູບແບບຕົວອັກສອນ (Font) 7 ຮູບແບບ ແຕ່ລະຮູບແບບຈະມີຕົວທໍາມະດາ, ຕົວໜາ ແລະ ຕົວເນີ້ງ, ເຊິ່ງລວມຮູບຕົວອັກສອນທັງໝົດແມ່ນ 1,911 ຮູບ ໃນນັ້ນໄດ້ແບ່ງຮູບຕົວອັກສອນອອກເປັນ 2 ສ່ວນຄື: 80% ແມ່ນສໍາລັບການຝຶກສອນ ແລະ 20% ແມ່ນສໍາລັບການທົດສອບ. ໃນຂັ້ນຕອນການຈື່ຈໍາຕົວອັກສອນແມ່ນໄດ້ນຳໃຊ້ແບບຈໍາລອງ ANN, CNN ແລະ KNN ໃນການຈື່ຈໍາ ແລະ ປຽບທຽບປະສິດທິພາບຂອງແຕ່ລະແບບຈໍາລອງ ໂດຍກໍານົດຮູບພາບ Input ຂະໜາດ 32x32 pixel. ຜົນຂອງການທົດລອງດ້ວຍການນຳໃຊ້ Confusion Matrix ມາວັດຝະສິດທິພາບຂອງແບບຈໍາລອງ ເຊິ່ງແບບຈໍາລອງຂອງ ANN ແມ່ນມີຄ່າຄວາມຖືກຕ້ອງ 96,60%, ຄ່າຂອງການຈຳແນກຜິດພາດແມ່ນ 3.40% ແລະ ແບບຈໍາລອງ CNN ແມ່ນມີຄວາມຖືກຕ້ອງເຖິງ 98.95%, ຄ່າຂອງການຈໍາແນກຜິດພາດແມ່ນ 1.05%. ສ່ວນແບບຈໍາລອງ KNN ແມ່ນມີຄວາມຖືກຕ້ອງ 79.06%, ຄ່າຂອງການຈໍາແນກຜິດພາດແມ່ນ 20.94% ເມື່ອສົມບຽບປະສິດທິພາບຂອງແບບຈໍາລອງທັງສາມແບບ ເຫັນວ່າແບບຈຳລອງ ANN ແມ່ນມີປະສິດທະພາບໃນການຈຳແນກຕົວອັກສອນໄດ້ດີກວ່າ KNN ສ່ວນ CNN ແມ່ນມີປະສິດທະພາບດີທີ່ສຸດໃນການຈໍາແນກຕົວອັກສອນໄຕທຳ ເຊິ່ງມີຄ່າຄວາມຖືກຕ້ອງເຖິງ 98.95%.

HiFi-GAN: ການສັງເຄາະສຽງໂດຍໃຊ້ເຄືອຂ່າຍແບບ Generative Adversarial Networks ສຳລັບການສັງເຄາະສຽງເວົ້າທີ່ມີປະສິດທິພາບ ແລະ ມີຄວາມຈັດເຈນສູງ (HiFi-GAN: Generative Adversarial Networks for Efficient and High Fidelity Speech Synthesis) ເປັນແບບຈຳລອງທີ່ຖືກພັດທະນາໂດຍ ບໍລິສັດ ກາກາວ ເອັນເຕີໄພຣ໌ (Kakao Enterprise) ເຊິ່ງພວກເຂົາໄດ້ສຶກສາກ່ຽວກັບບັນຫາຂອງ GAN (Generative Adversarial Networks) ໂດຍທີ່ພວກເຂົາຈະປັບປຸງປະສິດທິພາບການສຸ່ມຕົວຢ່າງ ແລະ ການໃຊ້ໜ່ວຍຄວາມຈຳ ເຊັ່ນວ່າ: ສຽງເວົ້າປະກອບດ້ວຍສັນຍານທີ່ເປັນຊ່ວງເວລາຕ່າງໆ ເຊິ່ງຈະໃຊ້ mean opinion score, MOS ເປັນຕົວປະເມີນປະສິດທິພາບແບບມະນຸດຂອງຊຸດຂໍ້ມູນນີ້ ສະແດງໃຫ້ເຫັນເຖິງຄວາມຄ້າຍຄືກັນລະຫວ່າງສຽງຂອງມະນຸດກັບແບບຈຳລອງ ໂດຍສຽງທີ່ໃຊ້ແມ່ນຢູ່ໃນຄື້ນ 22.05 kHz ເຊິ່ງໄວກວ່າແບບ Real-Time ເຖິງ 167.9 ເທົ່າເທິງ GPU Nvidia Tesla V100 ຕົວດຽວ. ພວກເຂົາຈົ້າໄດ້ໃຊ້ຊຸດຂໍ້ມູນ LJSpeech ເຊິ່ງມີຫຼາຍແບບຈໍາລອງການສັງເຄາະສຽງເວົ້າຖືກຝຶກອົບຮົມ, LJSpeech ປະກອບດ້ວຍ 13,100 ຄລິບສຽງສັ້ນຂອງຜູ້ເວົ້າຄົນດຽວທີ່ມີຄວາມຍາວທັງຫມົດປະມານ 24 ຊົ່ວໂມງ. ມີຄ່າ Mean Opinion Score (MOS (CI)) 4.05 (±0.08), ຄ່າຄວາມໄວເມື່ອປະມວນຜົນກັບ ຫນ່ວຍປະມວນຜົນຫລັກ(Speed on CPU (kHz)) 296.38 (×13.44), ຄ່າຄວາມໄວເມື່ອປະມວນຜົນກັບ ກາດສະແດງຜົນໜ້າຈໍ (Speed on GPU (kHz)) 26,169 (×1,186.80) ແລະ ຄ່າຜັນປ່ຽນ (# Param (M)) 1.46.

# ວິທີດຳເນີນການຄົ້ນຄວ້າ



## ວິທີສຶກສາ ແລະ ຄົ້ນຄວ້າ



### ການກຳນົດເນື້ອໃນ

ການຮຽນຮູ້ ແລະ ການຈື່ຈຳ ຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ ດ້ວຍການນຳໃຊ້ເຕັກນິກເຄືອຂ່າຍປະສາດທຽມ DNN ຂອງ Deep Learning ໂດຍນໍາໃຊ້ສະຖາປັດຕະຍະກໍາ RNN ດ້ວຍເຕັກນິກ YourTTS. ຊຶ່ງຂໍ້ມູນສຽງລວມທັງໝົດທີ່ນຳມາເປັນຊຸດຂໍ້ມູນມີທັງໝົດ 900 ປະໂຫຍກ ລວມເປັນຈຳນວນຄຳທີ່ບໍ່ຊ້ຳກັນຢູ່ປະມານ 600 ຄຳ ເຊິ່ງໄດ້ນຳເອົາແຕ່ລະປະໂຫຍກທີ່ເປັນຂໍ້ຄວາມ ຫຼື Text ມາຈາກບົດຄວາມທາງອິນເຕີເນັດ ແລະ ສື່ສິ່ງພິມຕ່າງໆ ມາສ້າງເປັນຊຸດຂໍ້ມູນສຽງສຳລັບການເຝິກສອນ ແລະ ທົດສອບແບບຈຳລອງ YourTTS.

YourTTS

RNN

DNN

ຮູບທີ 3.1 ຮູບພາບສະແດງພາບລວມຂອງເຕັກນິກເຄືອຂ່າຍປະສາດທຽມ DNN ຂອງ Deep Learning ໂດຍນໍາໃຊ້ສະຖາປັດຕະຍະກໍາ RNN ດ້ວຍເຕັກນິກ YourTTS

ຫຼັງຈາກເຮັດການເຝິກສອນ ແລະ ທົດສອບແບບຈຳລອງ YourTTS ແລ້ວ ນຳເອົາແບບຈຳລອງດັ່ງກ່າວໄປພັດທະນາເປັນ Web Site ເຊິ່ງຂັ້ນຕອນການສຶກສາແມ່ນໄດ້ແບ່ງອອກເປັນ 4 ຂັ້ນຕອນຄື:

#### ເກັບກຳຂໍ້ມູນສຽງມາຈາກບົດຄວາມທາງອິນເຕີເນັດ ແລະ ສື່ສິ່ງພິມຕ່າງໆທີ່ຢູ່ໃນຮູບແບບຂໍ້ຄວາມ ຫຼື Text ໂດຍນໍາເອົາ Text ດັ່ງກ່າວມາອັດສຽງເປັນຟາຍສຽງປະເພດ mp3. ຫຼັງຈາກນັ້ນ, ນຳໄປປັບປ່ຽນໄປເປັນຟາຍປະເພດ wav ເພື່ອນໍາໄປໃຊ້ໃນການເຝິກສອນ ແລະ ທົດສອບ.

#### ສຶກສາ ແລະ ນຳໃຊ້ແບບຈຳລອງຂອງ YourTTS ສຳລັບການຮຽນຮູ້ ແລະ ຈື່ຈຳ ຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ.

#### ປະເມີນປະສິດທິພາບຂອງແບບຈຳລອງ.

#### ນຳເອົາແບບຈຳລອງມາພັດທະນາເປັນ Web Site ທີ່ສາມາດຈື່ຈໍາຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ.

1

2

3

4

ການເກັບກຳ

ຂໍ້ມູນສຽງ

ສຶກສາ ແລະ ນຳໃຊ້ແບບຈຳລອງຂອງ YourTTS

ປະເມີນປະສິດທິພາບຂອງແບບຈຳລອງ

ນຳເອົາແບບຈຳລອງມາພັດທະນາເປັນ Website

ຮູບທີ 3.2 ຂັ້ນຕອນການສຶກສາ

### ການຄັດເລືອກພື້ນທີ່

ສຳລັບພື້ນທີ່ການເກັບກຳຂໍ້ມູນສຽງທີ່ນຳມາເປັນຊຸດຂໍ້ມູນໃນຄັ້ງນີ້ ແມ່ນໄດ້ມາຈາກບົດຄວາມທາງອິນເຕີເນັດ ແລະ ສື່ສິ່ງພິມຕ່າງໆທີ່ຢູ່ໃນຮູບແບບຂໍ້ຄວາມ ຫຼື Text ໂດຍນຳເອົາ Text ດັ່ງກ່າວມາອັດສຽງເປັນຟາຍສຽງປະເພດ mp3.

ສະຖານທີ່ການສຶກສາ ພວກຂ້າພະເຈົ້າ ໄດ້ດຳເນີນການຄົ້ນຄວ້າຢູ່ທີ່ ພາກວິຊາວິທະຍາສາດຄອມພິວເຕີ, ຄະນະວິທະຍາສາດທຳມະຊາດ, ມະຫາວິທະຍາໄລແຫ່ງຊາດ.

## ເຄື່ອງມືທີ່ໃຊ້ໃນການຄົ້ນຄວ້າ



#### ອຸປະກອນທາງດ້ານ Hardware

##### ກ. ເຄື່ອງທີ 1 ໃຊ້ໃນການສ້າງແບບຈໍາລອງ ແລະ Server:

###### ASUS ROG Strix G531GU-AL001T

###### Processor Intel® Core™ 17-9th Generation

###### Nvidia GeForce GTX 1660Ti Graphic Card

###### 24GB RAM

###### Operation: Microsoft Windows, Ubuntu

##### ຂ. ເຄື່ອງທີ 2 ໃຊ້ໃນການອອກແບບໜ້າຕາເວັບໄຊ (UI) ແລະ ກະກຽມເອກະສານ:

###### Lenovo Ideapad G40-70

###### Processor Intel® Core™ I5-4th Generation

###### Intel® HD Graphic Family Graphic Card

###### 8GB RAM

###### Operation: Microsoft Windows

ຄ. ເຄື່ອງທີ 3 ໃຊ້ໃນການເກັບກໍາຊຸດຂໍ້ມູນສຽງ

###### Oppo Reno 8T CPH2481

###### Mediatek MT8781 Helio G99 (6nm)

###### ColorOS 13, Android 13

###### 8GB RAM

#### ອຸປະກອນທາງດ້ານ Software

IDE: Visual Studio Code and Google Colab.

Framework and Library: TensorFlow, TF Lite, Coqui-AI, TTS, TensorBoard, NumPy, SciPy, Librosa, Torch (Torch, Torchaudio), Pandas, Trainer, Matplotlib and Flask.

Programming Languages: Python, HTML, CSS and JavaScript.

Programs Audio: Adobe Audition and Media Player.

Documentation: Microsoft Office 365

Server: Nginx

Web Browser: Microsoft Edge and Google Chrome

Dataset: Google Drive

## ການເກັບຮວບຮວມຂໍ້ມູນ













### ການຄັດເລືອກປະຊາກອນ

ການເກັບກຳຂໍ້ມູນໃນການສຶກສາຄັ້ງນີ້ ແມ່ນໄດ້ແຍກອອກເປັນ 3 ປະເພດຄື: ຂໍ້ມູນຕົວອັກສອນພາສາລາວ, ຂໍ້ມູນທີ່ເປັນຂໍ້ຄວາມ ຫຼື ປະໂຫຍກພາສາລາວ ແລະ ຂໍ້ມູນສຽງພາສາລາວ. ດັ່ງນີ້:



#### ຂໍ້ມູນຕົວອັກສອນພາສາລາວ

ຂໍ້ມູນຕົວອັກສອນພາສາລາວ ແມ່ນໄດ້ເລືອກເອົາຂໍ້ມູນຕົວອັກສອນພາສາລາວປະກອບມີ: ພະຍັນຊະນະ, ສະຫຼະ, ວັນນະຍຸດ ແລະ ສັນຍາລັກໃນພາສາລາວ. ດັ່ງທີ່ສະແດງໃນຕາຕະລາງ 3.1

ຕາຕະລາງທີ 3.1 ຕົວອັກສອນໃນພາສາລາວ

|  |  |  |
| --- | --- | --- |
| **ປະເພດຕົວອັກສອນ** | **ຕົວອັກສອນ** | **ຈຳນວນ** |
| ພະຍັນຊະນະ | ກ, ຂ, ຄ, ງ, ຈ, ສ, ຊ, ຍ, ດ, ຕ, ຖ, ທ, ນ, ຍ, ປ, ຜ, ຝ, ພ, ຟ, ມ, ຢ, ສ, ຮ, ວ, ຫູ, ອ, ຮ, ໜ(ຫນ), ໝ(ຫມ), ຫຼ(ຫລ) | 30 |
| ສະຫຼະ | xະ, xາ, xິ, xີ, xຶ, xື, xຸ, xູ, ເx, ເເx, xໍ, ໂx, ໄx,  ໃx, xົ, xັ, ຽ | 17 |
| ວັນນະຍຸດ | x່, x້, x໊, x໋ | 4 |
| ສັນຍາລັກອື່ນໆ | ໆ, x໌ | 2 |
| **ລວມ** | | **53** |

ຈາກຕາຕະລາງ 3.1 ເຫັນວ່າຕົວອັກສອນພາສາລາວປະກອບດ້ວຍ 4 ປະເພດໄດ້ແກ່: ພະຍັນຊະນະຈຳນວນ 30 ຕົວ, ສະຫຼະຈຳນວນ 17 ຕົວ, ວັນນະຍຸດຈໍານວນ 4 ຕົວ ແລະ ສັນຍາລັກອື່ນໆ 2 ຕົວ, ລວມທັງຕົວເລກ ແລະ ຕົວອັກສອນມີທັງໝົດ 53 ຕົວ.

#### ຂໍ້ມູນທີ່ເປັນຂໍ້ຄວາມ ຫຼື ປະໂຫຍກພາສາລາວ

ຂໍ້ມູນທີ່ເປັນຂໍ້ຄວາມ ຫຼື ປະໂຫຍກພາສາລາວ ແມ່ນໄດ້ເກັບກຳມາຈາກບົດຄວາມທາງອິນເຕີເນັດ ແລະ ສື່ສິ່ງພິມຕ່າງໆ. ດັ່ງທີ່ສະແດງໃນຕາຕະລາງ 3.2

ຕາຕະລາງທີ 3.2 ບົດຄວາມ ແລະ ສື່ສິ່ງພິມຕ່າງໆທາງອິນເຕີເນັດ

|  |  |  |  |
| --- | --- | --- | --- |
| ລຳດັບ | ຊື່ບົດຄວາມ ແລະ ສື່ສິ່ງພິມຕ່າງໆ | ຈຳນວນປະໂຫຍກ | ຈຳນວນຄຳທີ່ຊ້ຳກັນ |
| 1 | ມີທາງດີທາງງາມ ຕ້ອງຮູ້ຮັກສາບົວລະບັດ ໃຫ້ສົມກັບຄຳວ່າ: ລັດກັບປະຊາຊົນຊ່ວຍ ກັນເຮັດ ຊ່ວຍກັນສ້າງສາພັດທະນາໃຫ້ຈະເລີນ (ຜາຕຸງ, 2023) | 61 | 793 |
| 2 | ຄວນກຳຈັດຮື້ຖອນອອກແຕ່ຫົວທີ ກ່ອນຈະມີນ້ຳຖ້ວມອັ່ງຂັງໃນຕົວເມືອງ (ຜາຕຸງ, 2023) | 53 | 689 |
| 3 | ກ້າວເຂົ້າສູ່ລະດູຝົນແລ້ວ ບໍ່ຢາກໃຫ້ນ້ຳຖ້ວມຂັງ ຕ້ອງຈິງຈັງຮ່ວມມືກັນແກ້ໄຂ (ຜາຕຸງ, 2023) | 32 | 448 |
| 4 | ປູກພືດກິນໄດ້​ແທນຢາສູບ” ເວົ້າຄືຊິງ່າຍ ແຕ່ເຮັດແທ້ໆຍັງມີຫລາຍປັດໄຈທີ່ຕ້ອງແກ້ (ຜາຕຸງ, 2023) | 64 | 832 |
| 5 | ຫລາຍຄົນລະດົມກັນປູກຕົ້ນໄມ້ໃຫ້ເປັນປ່າ ແຕ່ຢ່າໃຫ້ໃຜຄົນໃດໜຶ່ງໃຊ້ອາຍາສິດໃນການທຳລາຍ (ທິດກໍ່າ, 2023) | 60 | 780 |
| 6 | ການຖິ້ມຂີ້ເຫຍື້ອຊະຊາຍ ເປັນບັນຫາທີ່ຍັງແພ່ຫລາຍໃນສັງຄົມລາວ (ພູຊ້າງນ້ອຍ, 2023) | 66 | 792 |
| 7 | ເອົາໃຈໃສ່ກະກຽມໃຫ້ດີ ແລະ ສອບເສັງໃຫ້ໄດ້ຕາມຄາດໝາຍຈຶ່ງບໍ່ຕົກໃສ່ຄຳວ່າ ເຮັດນາໝົດປີໄຟໄໝ້ເລົ້າເຂົ້າ (ຜາຕຸງ, 2023) | 62 | 744 |
| 8 | 4 ທ່າແຮງໃນການພັດທະເສດຖະກິດ ນຳເອົາປະຊາຊົນເມືອງຊຳເໜືອ ແຂວງຫົວພັນໃຫ້ຫລຸດພົ້ນອອກຈາກຄວາມທຸກຍາກ (ຜາຕຸງ, 2023) | 71 | 923 |
| 9 | ການຂຸດຄົ້ນແຮ່ທາດ ທີ່ບໍ່ໄປຕາມມາດຕະຖານ ແລະ ກົດໝາຍ ກາຍເປັນບັນຫາເຄັ່ງຮ້ອນຂອງຊາດ (ມຶກບໍ່ແຫ້ງ, 2023) | 74 | 935 |
| 10 | ຄວນຄຳນຶງເຖິງຄວາມປອດໄພກ່ອນຈະກິນອາຫານທີ່ປຸງແຕ່ງດ້ວຍເຫັດປ່າ (ຜາຕຸງ, 2023) | 76 | 958 |
| 11 | ກູ້ເງິນອອນລາຍງ່າຍ ບໍ່ມີຈິງ ລະວັງມິດສາຊີບລໍ້ລວງເອົາຂໍ້ມູນສ່ວນຕົວ (ຜາຕຸງ, 2023) | 71 | 880 |
| 12 | ບານເຕະທີມຊາດລາວໄປບໍ່ເຖິງ ຕາມທີ່ປວງຊົນລາວຕັ້ງເປົ້າໝາຍ ເຊິ່ງມີຫລາຍຢ່າງຕ້ອງພັດທະນາ (ທິດກໍ່າ, 2023) | 63 | 769 |
| 13 | ທາງການສັ່ງປິດຮ້ານແລກປ່ຽນເງິນຕາແລ້ວກໍຈິງ ແຕ່ຍັງປາກົດມີການລັກແລກປ່ຽນນອກລະບົບແບບຜິດກົດໝາຍຢູ່ (ທິດກໍ່າ, 2023) | 72 | 893 |
| 14 | ຍຸກຂໍ້ມູນຂ່າວສານທັນສະໄໝ ຢ່າໃຫ້ຕົກເປັນເຫຍື່ອກຸ່ມຕົ້ມຕຸນຂາຍຝັນລໍ້ລວງໃຫ້ເກີດຄວາມໂລບ (ທິດກໍ່າ, 2023) | 75 | 930 |
| ລວມ | | 900 | 11,366 |

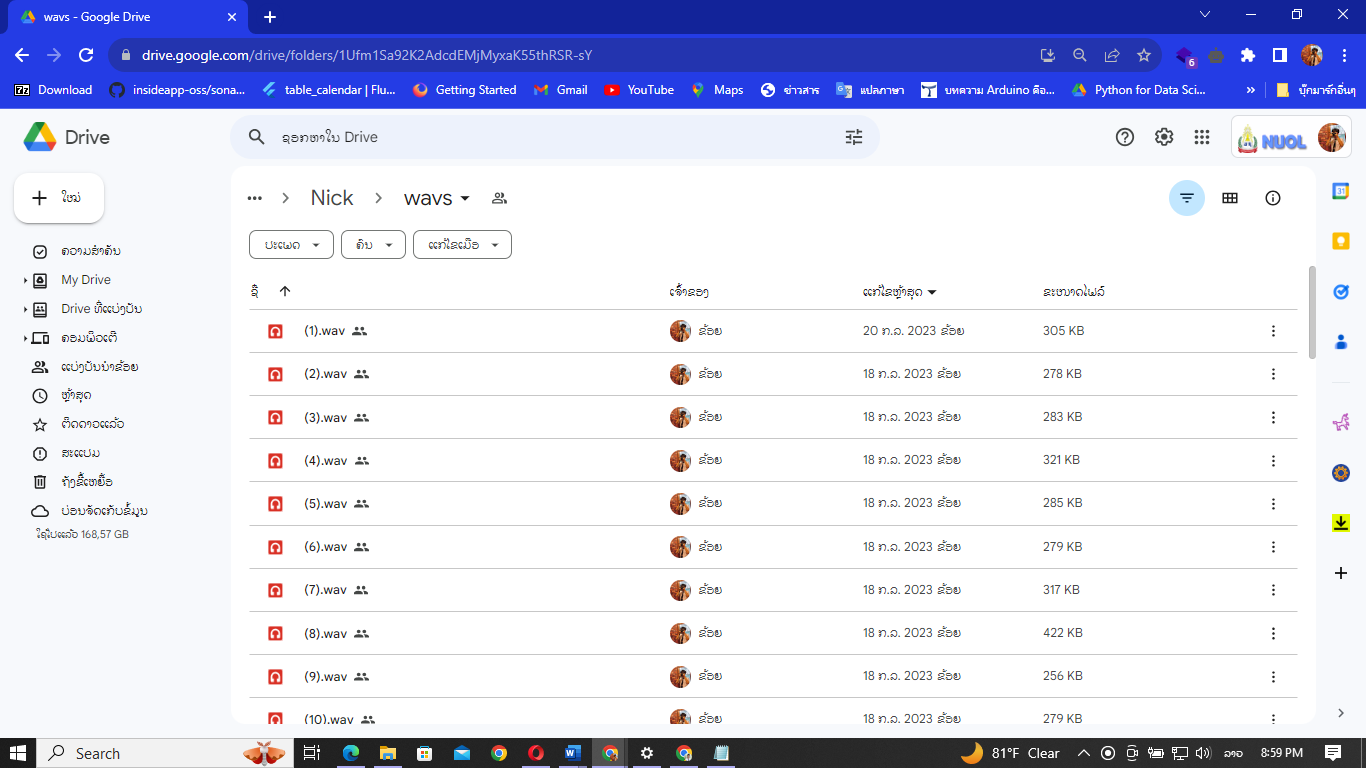
ຈາກຕາຕະລາງ 3.2 ເຫັນວ່າປະກອບມີບົດຄວາມ ແລະ ສື່ສິ່ງພິມຕ່າງໆທັງໝົດຈຳນວນ 14 ລາຍການ, ຈຳນວນປະໂຫຍກທີ່ແຍກຕາມການບັນທຶກສຽງເປັນຟາຍປະເພດ mp3 ມີທັງໝົດ 900 ປະໂຫຍກ, ຈຳນວນຄຳທີ່ຊ້ຳກັນມີທັງໝົດ 11,366 ຄຳ ແລະ ຈຳນວນຄຳທີ່ບໍ່ຊ້ຳກັນມີທັງໝົດ 600 ຄຳ.

#### ຂໍ້ມູນສຽງພາສາລາວ

ຂໍ້ມູນສຽງພາສາລາວ ແມ່ນໄດ້ເກັບກຳມາຈາກການບັນທຶກສຽງຜ່ານທາງໂທລະສັບມືຖື (Android OS) ລວມຈຳນວນສຽງທີ່ບັນທຶກມີທັງໝົດເປັນຟາຍສຽງປະເພດ mp3 ຈໍານວນ 900 ຟາຍສຽງ. ຫຼັງຈາກນັ້ນ, ນຳຟາຍສຽງປະເພດ mp3 ໄປປ່ຽນເປັນຟາຍສຽງປະເພດ wav ພ້ອມທັງຕັດເອົາສະເພາະສຽງທີ່ເຮົາບັນທຶກ (ຫຼື ຕັດຊ່ອງວ່າງຂອງສຽງທີ່ບັນທຶກ) ດ້ວຍໄລບຣາລີ Librosa ແລະ ຕັດສຽງລົບກວນອ້ອມຂ້າງດ້ວຍໂປຣແກຣມ Adobe Audition.

### ການກຳນົດຈຳນວນຕົວຢ່າງປະຊາກອນ

ຈຳນວນປະຊາກອນຊຸດຂໍ້ມູນ (Dataset) ໄດ້ດາວໂຫຼດບົດຄວາມ ແລະ ຂ່າວສານຈາກສື່ສິ່ງພິມຕ່າງໆໃນອິນເຕີເນັດເປັນຈຳນວນຄຳທີ່ຊ້ຳກັນ 11,366 ຄຳ ຕົກເປັນຈຳນວນຄຳທີ່ບໍ່ຊ້ຳກັນມີທັງໝົດ 600 ຄຳ ແລະ ໄດ້ສ້າງຊຸດຂໍ້ມູນສຽງຂຶ້ນເອງໂດຍນຳໃຊ້ໂທລະສັບມືຖື (Android OS) ຜ່ານແອັບພິເຄເຊີນ Voice Recorder ຂອງ Google Play ຈຳນວນ 900 ຟາຍ. ຫຼັງຈາກນັ້ນ, ນຳຟາຍສຽງທັງໝົດໄປຈັດເກັບໄວ້ໃນ Google Drive ດັ່ງທີ່ສະແດງໃນຮູບທີ 3.3



ຮູບທີ 3.3 ຮູບພາບສະແດງຂະບວນການເກັບກໍາຊຸດຂໍ້ມູນໃນ Google Drive

### ວິທີສຸ່ມຕົວຢ່າງປະຊາກອນ

ຈຳນວນຊຸດຂໍ້ມູນຂອງຟາຍສຽງພາສາລາວມີທັງໝົດ 900 ຟາຍສຽງໄດ້ແບ່ງອອກເປັນ 2 ຊຸດຄື: ຊຸດຂໍ້ມູນທີ່ໃຊ້ໃນການເຝີກລະບົບ (Training set) ຈະແບ່ງເປັນ 95% ຈຳນວນ 855 ຟາຍສຽງ ແລະ ຊຸດຂໍ້ມູນທີ່ໃຊ້ໃນການທົດສອບ (Test set) ຈະແບ່ງເປັນ 5% ຈຳນວນ 45 ຟາຍສຽງ ໂດຍໄດ້ນຳໃຊ້ຫຼັກການຂອງການຮຽນຮູ້ຂັ້ນສູງ (Deep Learning) ດັ່ງນີ້:

Dataset = D = DTrain + DTest

D ແມ່ນຊຸດຂໍ້ມູນທັງໝົດ

DTrain  ແມ່ນຊຸດຂໍ້ມູນທີ່ໄວ້ໃຊ້ໃນການຮຽນຮູ້ເພື່ອສ້າງແບບຈຳລອງ

DTest ແມ່ນຊຸດຂໍ້ມູນທີ່ໄວ້ໃຊ້ໃນການທົດສອບແບບຈຳລອງ

metadata.csv

100%

metadata\_train.csv

95%

metadata\_eval.csv

5%

Train

Model

Accuracy

ຮູບທີ 3.4 ຮູບພາບສະແດງໃຫ້ເຫັນເຖິງການແບ່ງຂໍ້ມູນເພື່ອການ Train ແລະ Test

ຈາກຮູບທີ 3.4 ຈະເຫັນຟາຍຊື່ metadata.csv ແມ່ນເປັນຕົວເກັບກຳຊຸດຂໍ້ມູນທັງໝົດ ໂດຍຟາຍທີ່ໃຊ້ເຝິກແບບຈຳລອງມີຊື່ວ່າ: metadata\_train.csv ແລະ metadata\_eval.csv ເປັນຊຸດຂໍ້ມູນທີ່ໃຊ້ທົດສອບແບບຈຳລອງ.

### ບັນດາຂໍ້ມູນການສຶກສາ

ບັນດາຂໍ້ມູນການສຶກສາຂອງຕົວອັກສອນຕາມຫຼັກໄວຍາກອນລາວແລ້ວ ພາສາລາວເປັນພາສາທີ່ຂຽນແຕ່ຊ້າຍຫາຂວາ ໂດຍໄດ້ແບ່ງເສັ້ນຕາມລະດັບແນວນອນແບ່ງອອກເປັນ 5 ລະດັບ (ດັ່ງທີ່ສະແດງໃນຮູບທີ 3.5) ຄື:

#### ລະດັບທີ່ມີສຽງເທີງ 2 (Upper Level 2) ໄດ້ແກ່ ວັນນະຍຸດ ແລະ ໄມ້ກະລັນ ດັ່ງນີ້: x່, x້, x໊, x໋

#### ລະດັບທີ່ມີສຽງເທີງ 1 (Upper Level 1) ໄດ້ແກ່ ສະຫຼະທາງເທິງ ດັ່ງນີ້: xິ, xີ, xຶ, xື xໍ, xົ, xັ

#### ລະດັບທີ່ມີສຽງກາງ (Centre Level) ໄດ້ແກ່ ພະຍັນຊະນະ, ສະຫຼະທາງໜ້າ, ສະຫຼະທາງຫຼັງ ແລະ ເຄື່ອງໝາຍພິເສດ ເຊິ່ງສາມາດແບ່ງອອກເປັນ 3 ໝວດຄື:

###### ພະຍັນຊະນະທີ່ມີສຽງຄວາມສູງທຳມະດາຄື: ກ,​ ຂ, ຄ, ຈ, ສ, ຍ, ດ, ຕ, ທ, ນ, ບ, ຜ, ຟ, ມ, ລ, ຣ, ວ, ຫ, ອ, ຮ

###### ພະຍັນຊະນະທີ່ມີສຽງສູງກາຍໄປຫາລະດັບເທິງ 1 ຄື: ປ, ຝ, ຟ, ຢ

###### ພະຍັນຊະນະທີ່ມີສຽງຍາວລົງໄປຫາລະດັບລຸ່ມ: ງ, ຊ, ຖ, ໆ, ຽ

#### ລະດັບທີ່ມີສຽງລຸ່ມ 1 (Lower Level 1) ໄດ້ແກ່ພະຍັນຊະ “xຼ”

#### ລະດັບທີ່ມີສຽງລຸ່ມ 2 (Lower Level 2) ໄດ້ແກ່ ສະຫຼະທາງລຸ່ມຄື: xຸ, xູ

ສິນຄ້ານີ້ຫຼຸດລາຄາ

Upper Level 1

Upper Level 2

Center Level

Lower Level 1

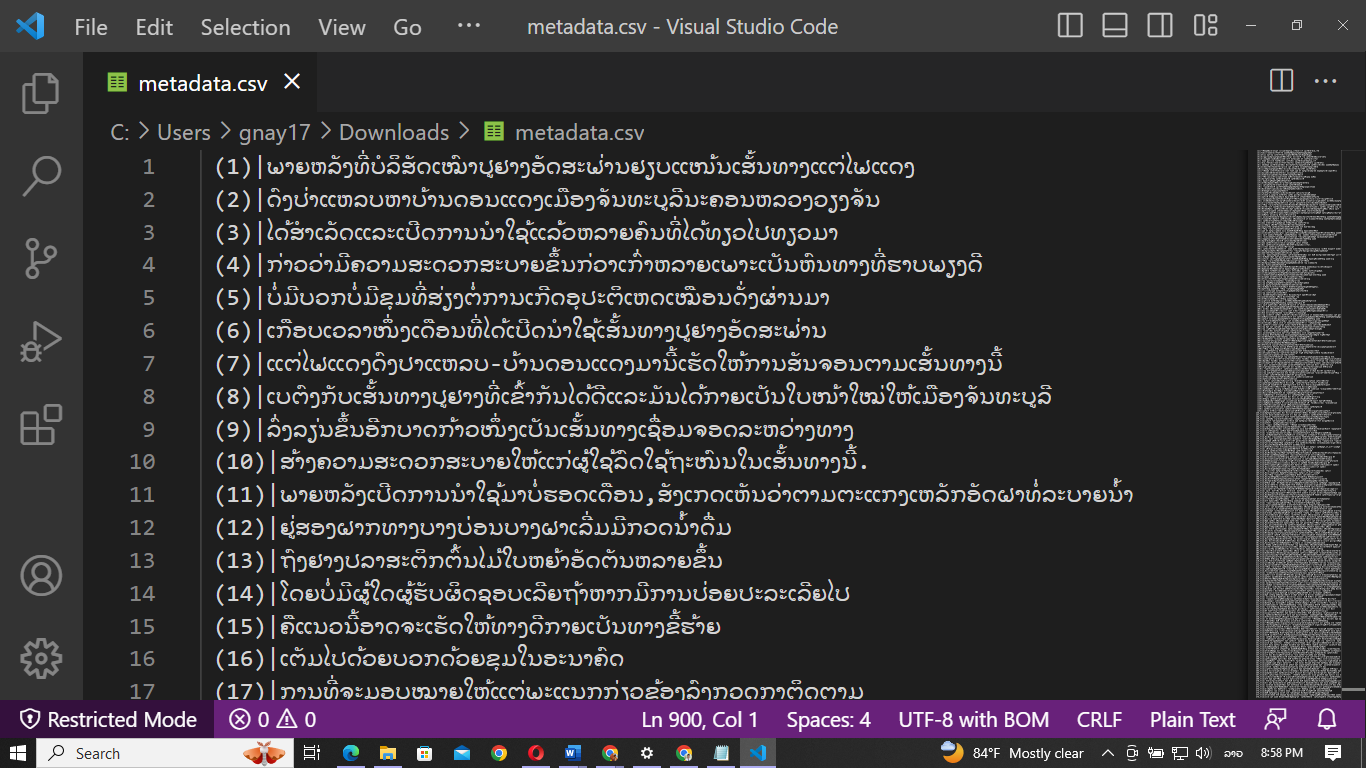
Lower Level 2

ຮູບທີ 3.5 ຮູບພາບສະແດງລະດັບຂອງຕົວອັກສອນ

ຈາກຮູບທີ 3.5 ສະແດງໃຫ້ເຫັນວ່າພາສາລາວແມ່ນໄດ້ແບ່ງລະດັບທີ່ມີສຽງຕົວອັກສອນອອກເປັນ 5 ລະດັບຄື: ລະດັບເທີງ 2 (Upper Level 2), ລະດັບເທີງ 1 (Upper Level 1), ລະດັບກາງ (Centre Level), ລະດັບລຸ່ມ 1 (Lower Level 1) ແລະ ລະດັບລຸ່ມ 2 (Lower Level 2).

### ວິທີເກັບກຳຂໍ້ມູນ

ວິທີເກັບກຳຂໍ້ມູນໃນຄັ້ງນີ້ ໄດ້ດຳເນີນການດາວໂຫຼດບົດຄວາມ ແລະ ສື່ສິ່ງພິມຕ່າງໆ ແລ້ວນຳເນື້ອໃນບົດຄວາມ ແລະ ສື່ສຶ່ງພິມຕ່າໆ ມາອັດສຽງເປັນຟາຍປະເພດ mp3. ຈາກນັ້ນ, ນຳເອົາຟາຍ mp3 ທີ່ໄດ້ເຮັດການບັນທຶກສຽງທັງໝົດໄປປ່ຽນເປັນຟາຍສຽງປະເພດ wav ພ້ອມທັງຕັດເອົາສະເພາະສຽງທີ່ເຮົາບັນທຶກ (ຫຼື ຕັດຊ່ອງວ່າງຂອງສຽງທີ່ບັນທຶກ) ດ້ວຍໄລບຣາລີ Librosa ແລະ ຕັດສຽງລົບກວນອ້ອມຂ້າງດ້ວຍໂປຣແກຣມ Adobe Audition. ຫຼັງຈາກນັ້ນ, ນໍາເອົາຊື່ຂອງຟາຍ wav ແລະ ປະໂຫຍກພາສາລາວທີ່ໃຊ້ອັດສຽງມານັ້ນ ໄປຈັດເກັບເປັນຟາຍ csv (ຊື່ຟາຍ: metadata.csv) ແລະ ໄດ້ເຮັດການແບ່ງຟາຍ metadata.csv ອອກເປັນ 2 ຟາຍຄື: ຟາຍທີ່ໃຊ້ໄວ້ໃນການເຝິກສອນແບບຈຳລອງ (ຊື່ຟາຍ: metadata\_train.csv) ແລະ ຟາຍທີ່ໃຊ້ໄວ້ໃນການທົດສອບແບບຈຳລອງ (ຊື່ຟາຍ: metadata\_eval.csv) ໂດຍການແບ່ງຟາຍແມ່ນໄດ້ອີງຕາມກົດ 95-5 ທີ່ໄດ້ກຳນົດໄວ້ກ່ອນໜ້ານີ້ (ດັ່ງຫົວຂໍ້ [3.3.3](#_ວິທີສຸ່ມຕົວຢ່າງປະຊາກອນ)) ໂດຍນໍາໃຊ້ໂປຣແກມ Visual Studio Code, ດັ່ງທີ່ສະແດງໃນຮູບທີ 3.6 ລຸ່ມນີ້:



ຮູບທີ 3.6 ຮູບພາບສະແດງຊຸດຂໍ້ມູນທັງໝົດໃນຟາຍ metadata.csv

ຈາກຮູບທີ 3.6 ຈະເຫັນວ່າມີ 2 ພາກສ່ວນຄື: ພາກສ່ວນທີ່ເປັນຊື່ຟາຍ wav ແລະ ພາກສ່ວນທີ່ເປັນປະໂຫຍກພາສາລາວທີ່ໃຊ້ອັດສຽງ ແລະ ມີເຄື່ອງໝາຍຂັ້ນກາງຄື Pipe ( | ) ເພື່ອໃຫ້ແບບຈຳລອງສາມາດເຂົ້າໃຈຂໍ້ຄວາມຈາກສຽງທີ່ອັດໄວ້ແຕ່ລະຄຳ.

## ການວິເຄາະຂໍ້ມູນ



### ຂະບວນການວິເຄາະຂໍ້ມູນ

ຂະບວນການວິເຄາະຂໍ້ມູນແມ່ນໄດ້ອີງໃສ່ຫຼັກການ ແລະ ທິດສະດີຂອງ ການຮຽນຮູ້ຂັ້ນສູງ (Deep learning) ແລະ ການປະມວນຜົນຂໍ້ຄວາມສຽງ (YourTTS), ເຊິ່ງການປະມວນຜົນ ແລະ ສະແດງຜົນຕ່າງໆ ແມ່ນນໍາໃຊ້ພາສາ Python ເປັນພາສາຫຼັກໃນການຂຽນໂປຣແກຣມ. ສະນັ້ນ, ຈິ່ງໄດ້ກຳນົດຂັ້ນຕອນໄວ້ດັ່ງລຸ່ມນີ້:

#### ການຈັດກຽມຂໍ້ມູນສຽງ

ຊຸດຂໍ້ມູນສຽງແມ່ນໄດ້ກຽມຢູ່ 3 ຊຸດຄື: ຊຸດຂໍ້ມູນສຽງທີ່ກຽມໄວ້ເພື່ອໃຫ້ແບບຈຳລອງຈື່ຈຳຂໍ້ຄວາມຂຽນ ແລະ ສຽງພາສາລາວ (ກໍຄືຟາຍ metadata.csv, ດັ່ງຫົວຂໍ້ [3.3.5](#_ເຄື່ອງມືທີ່ໃຊ້ໃນການຄົ້ນຄວ້າ)), ຊຸດຂໍ້ມູນສຽງທີ່ກຽມໄວ້ເພື່ອເຝິກສອນແບບຈຳລອງ (ກໍຄືຟາຍ metadata\_train.csv, ດັ່ງຫົວຂໍ້ [3.3.5](#_ເຄື່ອງມືທີ່ໃຊ້ໃນການຄົ້ນຄວ້າ)) ແລະ ຊຸດຂໍ້ມູນສຽງທີ່ກຽມໄວ້ເພື່ອທົດສອບແບບຈຳລອງ (ກໍຄືຟາຍ metadata\_eval.csv, ດັ່ງຫົວຂໍ້ [3.3.5](#_ເຄື່ອງມືທີ່ໃຊ້ໃນການຄົ້ນຄວ້າ)).

#### ການປະມວນຜົນຟາຍສຽງຂັ້ນຕົ້ນ

ກ່ອນທີ່ຈະເຂົ້າສູ່ຂະບວນການປະມວນຜົນ ຕ້ອງໄດ້ກະກຽມດັ່ງນີ້:

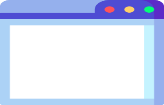
###### ກຳຈັດສຽງລົບກວນອ້ອມຂ້າງ (Eliminate Noise)

###### ຕັດເອົາສະເພາະສຽງທີ່ເຮົາບັນທຶກ (ຫຼື ຕັດຊ່ອງວ່າງຂອງສຽງທີ່ບັນທຶກ)

###### ການເຮັດໃຫ້ສຽງເປັນປົກກະຕິ

#### ການນຳໃຊ້ແບບຈຳລອງ YourTTS ແລະ ການປະມວນຜົນ

ການນຳໃຊ້ແບບຈຳລອງ YourTTS ແມ່ນໄດ້ນໍາເອົາຊຸດຂໍ້ມູນຟາຍສຽງທີ່ໄດ້ມາມີທັງຫມົດ 900 ຟາຍສຽງ (ເປັນຟາຍ .wav) ນໍາໄປປະມວນຜົນ, ເຝິກສອນ ແລະ ທົດສອບແບບຈໍາລອງໃນ Google Colab ໂດຍມີຂັ້ນຕອນການປະມວນຜົນດັ່ງທີ່ສະແດງໃນຮູບທີ 3.7 ລຸ່ມນີ້:



**root(vits\_tts)**

**mailabs**

**vits\_vctk\_datetime**

**lo-la**

**by\_book**

**female**

**Nick**

**wavs**

**metadata.csv**

**metadata\_train.csv**

**metadata\_** **eval.csv**

**config\_se.json**

**model\_se.pth**

**train\_vits\_tts.py**

**Google Colab**

**Main Folder**

**Folder**

**Sub Folder**

**File**

Request

Request Coding

Request path folder wavs and csv file

ຮູບທີ 3.7 ຮູບພາບສະແດງຂັ້ນຕອນການປະມວນຜົນແບບຈຳລອງ YourTTS

ຈາກຮູບທີ 3.7 ສະແດງໃຫ້ເຫັນຂັ້ນຕອນການປະມວນຜົນແບບຈໍາລອງ YourTTS ໃນ Google Colab ເຊິ່ງຈະມີໂຟເດີຫຼັກ (Main Folder) ເປັນຊື່ root ຂອງແບບຈໍາລອງ (root(vits\_tts)) ແລະ ໃນໂຟເດີຫຼັກຈະມີໂຟເດີ (Folder) ຈໍານວນ 6 ໂຟເດີ ແລະ ຟາຍ (File) ຈໍານວນ 6 ຟາຍ ໃນນັ້ນຈະມີໂຟເດີທີ່ມານຳແບບຈໍາລອງ YourTTS ຢູ່ຈໍານວນ 1 ໂຟເດີ ແລະ ຟາຍຈຳນວນ 3 ຟາຍ. ສ່ວນທີ່ເຫຼືອເປັນໂຟເດີ ແລະ ຟາຍທີ່ພວກຂ້າເຈົ້າໄດ້ສ້າງຂຶ້ນເພື່ອໃຊ້ໃນການປະມວນຜົນ, ເຝິກສອນ ແລະ ທົດສອບແບບຈຳລອງ ມີດັ່ງນີ້:

1. ໂຟເດີຫຼັກ root(vits\_tts) ເປັນຊື່ root ຂອງແບບຈຳລອງ YourTTS
2. ໂຟເດີ mailabs ແມ່ນໂຟເດີທີ່ມານໍາແບບຈໍາລອງ YourTTS ເປັນໂຟເດີທີ່ໃຊ້ໄວ້ເກັບຂໍ້ມູນຕ່າໆເພື່ອປະມວນຜົນ
3. ຟາຍ train\_vits\_tts.py ແມ່ນຟາຍທີ່ມານຳແບບຈຳລອງ YourTTS ເປັນຟາຍທີ່ໃຊ້ເພື່ອປະມວນຜົນ, ເຝິກສອນ, ທົດສອບແບບຈຳລອງ ແລະ ສ້າງແບບຈຳລອງ ສາມາດປ່ຽນຄ່າຕ່າງໆໃນຟາຍນີ້ໄດ້ເຊັ່ນ:​ ປ່ຽນຊື່ path ໃນການເອີ້ນໃຊ້ຊຸດຂໍ້ມູນພາສາລາວເປັນຕົ້ນ
4. ຟາຍ config\_se.json ແມ່ນຟາຍທີ່ມານໍາແບບຈໍາລອງ YourTTS ເປັນຟາຍແບບຢ່າງທີ່ໃຊ້ໄວ້ບັນທຶກຂໍ້ມູນຕ່າງໆທີ່ໄດ້ຈາກຟາຍ train\_vits\_tts.py ເພື່ອນຳໄປສະແດງຜົນໄດ້ຮັບໃນການປະມວນຜົນ, ການເຝິກສອນ ແລະ ການທົດສອບແບບຈຳລອງ
5. ຟາຍ model\_se.pth ແມ່ນຟາຍທີ່ມານໍາແບບຈໍາລອງ YourTTS ເປັນຟາຍແບບຢ່າງທີ່ໃຊ້ໄວ້ບັນທຶກຂໍ້ມູນຕ່າງໆທີ່ໄດ້ຈາກຟາຍ train\_vits\_tts.py ອອກໄປເປັນຟາຍໂມເດວທີ່ປະມວນຜົນສຳເລັດສົມບູນແລ້ວ
6. ໂຟເດີ vits\_vctk\_datetime ແມ່ນໂຟເດີທີ່ໄດ້ຈາກການປະມວນຜົນ, ການເຝິກສອນ, ການທົດສອບ ແລະ ການສ້າງແບບຈຳລອງສຳເລັດສົມບູນແລ້ວ ເປັນໂຟເດີທີ່ເກັບຂໍ້ມູນຕ່າງໆເປັນຕົ້ນແມ່ນຟາຍ checkpoint\_100000.pth, speakers.pth ແລະ config.json ເພື່ອນໍາໄປໃຊ້ງານໃນ Website
7. ໂຟເດີ lo-la, by\_book, female ແລະ Nick ແມ່ນໂຟເດີທີ່ພວກຂ້າພະເຈົ້າໄດ້ສ້າງຂຶ້ນເພື່ອໄວ້ໃຊ້ໃນການປະມວນຜົນ
8. ໂຟເດີ wavs ແມ່ນໂຟເດີທີ່ໃຊ້ເກັບຊຸດຂໍ້ມູນສຽງ (Dataset) ທີ່ຜ່ານການປະມວນຜົນຟາຍສຽງຂັ້ນຕົ້ນແລ້ວ
9. ຟາຍ metadata.csv ແມ່ນຟາຍທີ່ພວກຂ້າເຈົ້າໄດ້ສ້າງຂຶ້ນເພື່ອໃຫ້ແບບຈຳລອງຈື່ຈຳຂໍ້ຄວາມຂຽນ ແລະ ສຽງພາສາລາວ
10. ຟາຍ metadata\_train.csv ແມ່ນຟາຍທີ່ພວກຂ້າເຈົ້າໄດ້ສ້າງຂຶ້ນເພື່ອເຝິກສອນແບບຈຳລອງ
11. ຟາຍ metadata\_ eval.csv ແມ່ນຟາຍທີ່ພວກຂ້າເຈົ້າໄດ້ສ້າງຂຶ້ນເພື່ອທົດສອບແບບຈຳລອງ

#### ການປະເມີນປະສິດທິພາບແບບຈຳລອງ

ການປະເມີນປະສິດທິພາບແບບຈຳລອງ ແມ່ນໄດນຳໃຊ້ການປະເມີນຢູ່ 3 ແບບຄື: ການປະເມີນປະສິດທິພາບການປະມວນຜົນໂດຍລວມຂອງແບບຈຳລອງ (TrainItorstats), ການປະເມີນປະສິດທິພາບຈາກການທົດສອບແບບຈຳລອງ (EvalStats) ແລະ ການປະເມີນປະສິດທິພາບຈາກການເຝິກສອນແບບຈຳລອງແບບເປັນຮອບ (TrainEpochStats).

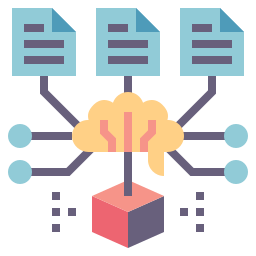
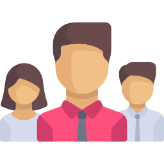
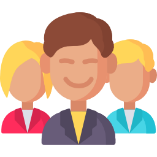
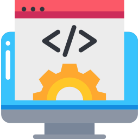
### ການນຳແບບຈຳລອງໄປໃຊ້ງານ

ການນຳແບບຈຳລອງໄປໃຊ້ງານ ພວກຂ້າພະເຈົ້າໄດ້ສ້າງເວັບໄຊຂຶ້ນມາໂດຍນຳໃຊ້ Flask Framework ໃນການພັດທະນາເວັບໄຊເພື່ອການເອີ້ນໃຊ້ງານແບບຈໍາລອງທີ່ໄດ້ບັນທຶກອອກເປັນຟາຍ checkpoint\_100000.pth, speakers.pth ແລະ config.json, ພາສາໂປຣແກຣມທີ່ໄດ້ນໍາໃຊ້ໃນການສ້າງເວັບໄຊຫຼັກໆແມ່ນ Python ໃຊ້ເປັນໂຕຄວບຄຸມການເຮັດວຽກ ແລະ ປະມວນຜົນ (ສ່ວນເວີເຊີນທີ່ໃຊ້ແມ່ນ Python 3.10), HTML ໃຊ້ເພື່ອສະແດງຂໍ້ມູນ ຫຼື ຂໍ້ຄວາມຕ່າງໆ, CSS ແລະ JavaScript ໃຊ້ເພື່ອປັບຄວາມສວຍງາມຂອງເວັບໄຊ. ສຳລັບເຄື່ອງມືທີ່ໃຊ້ຂຽນເວັບໄຊແມ່ນ Visual Studio Code.

## ການວິເຄາະລະບົບ



### ແຜນວາດລວມຂອງລະບົບ



**ຜູ້ໃຊ້ງານ**

**ຜູ້ໃຊ້ງານ**

**Google Cloud Speech API**

**ແບບຈຳລອງ**

ປ້ອນຂໍ້ຄວາມ

ສຽງອ່ານຂໍ້ຄວາມ

ຂໍ້ຄວາມທີ່ປ້ອນ

ສຽງອ່ານຂໍ້ຄວາມ

ຂໍ້ຄວາມ

ປ້ອນຟາຍສຽງ

ຂໍ້ຄວາມ

ຟາຍສຽງທີ່ປ້ອນ

ແຜນວາດທີ 3.1 ແຜນວາດສະແດງການໃຊ້ງານ Website

### ແຜນວາດ Use Case Diagram

#### ຄົ້ນຫາ Actor

ເວັບໄຊການແປຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ ແລະ ການແປຂໍ້ຄວາມສຽງພາສາລາວໄປເປັນຂໍ້ຄວາມຂຽນ ຂອງພວກຂ້າພະເຈົ້າແມ່ນປະກອບມີ 1 Actor ຄື: ຜູ້ໃຊ້ງານເວັບໄຊ (User).

#### ຄົ້ນຫາ Use Case ທີ່ໂຕ້ຕອບ Actor ໂດຍກົງ

ສາມາດຄົ້ນຫາ Use Case ທີ່ໂຕ້ຕອບ ຜູ້ໃຊ້ງານເວັບໄຊ (User) ໄດ້ດັ່ງນີ້:

* Text -To-Speech (TTS)

###### Input Text

###### Speech as Steaming

* Speech-To-Text (STT)

###### Speech

###### Text and Streaming

#### Use Case Diagram

ຜ່ານການວິເຄາະຫາ Actor ແລະ Use Case ຈາກຂ້າງເທິງນັ້ນແລ້ວ ເຮົາກໍໄດ້ Use Case Diagram ໃນລະດັບ System Context Model, ເຊິ່ງມີຄວາມສໍາພັນລະຫວ່າງ Use Case ແລະ Actor ດັ່ງລຸ່ມນີ້:

**Lao Text-To-Speech and Lao Speech-To-Text**

Input Text

TTS

Speech as Steaming

<< include >>

<<include>>

STT

<<include>>

Text and Streaming

<< include >>

User

Speech

ແຜນວາດທີ 3.2 ແຜນວາດສະແດງ Use Case Diagram ຂອງ Website ການແປຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ ແລະ ການແປຂໍ້ຄວາມສຽງໄປເປັນຂໍ້ຄວາມຂຽນ

#### ຄໍາອະທິບາຍ Use Case

ຕາຕະລາງທີ 3.3 ຄໍາອະທິບາຍ Text-To-Speech

|  |  |
| --- | --- |
| Use Case Title: Text-To-Speech | Use Case ID: 1 |
| Primary Actor: User | |
| **Main Flow:** ໃນກໍລະນີທີ່ User ຕ້ອງການໃຊ້ Text-To-Speech, User ຕ້ອງປ້ອນ: ຂໍ້ຄວາມ ເພື່ອໃຊ້ງານການແປງຂໍ້ຄວາມໄປເປັນສຽງ | |
| **Exceptional Flow 1:** ໃນກໍລະນີທີ່ User ປ້ອນຂໍ້ມູນບໍ່ຄົບຖ້ວນ, ລະບົບຈະແຈ້ງເຕືອນວ່າ “ກະລຸນາປ້ອນຂໍ້ມູນໃຫ້ຄົບຖ້ວນ”. | |

ຈາກ Main Flow ຂອງ Use Case Text-To-Speech ແມ່ນບໍ່ສາມາດຄົ້ນຫາ Class ແລະ Attribute ໄດ້ເນື່ອງຈາກວ່າບ່ໄດ້ຈັດເກັບຂໍ້ມູນທີ່ປ້ອນໄວ້.

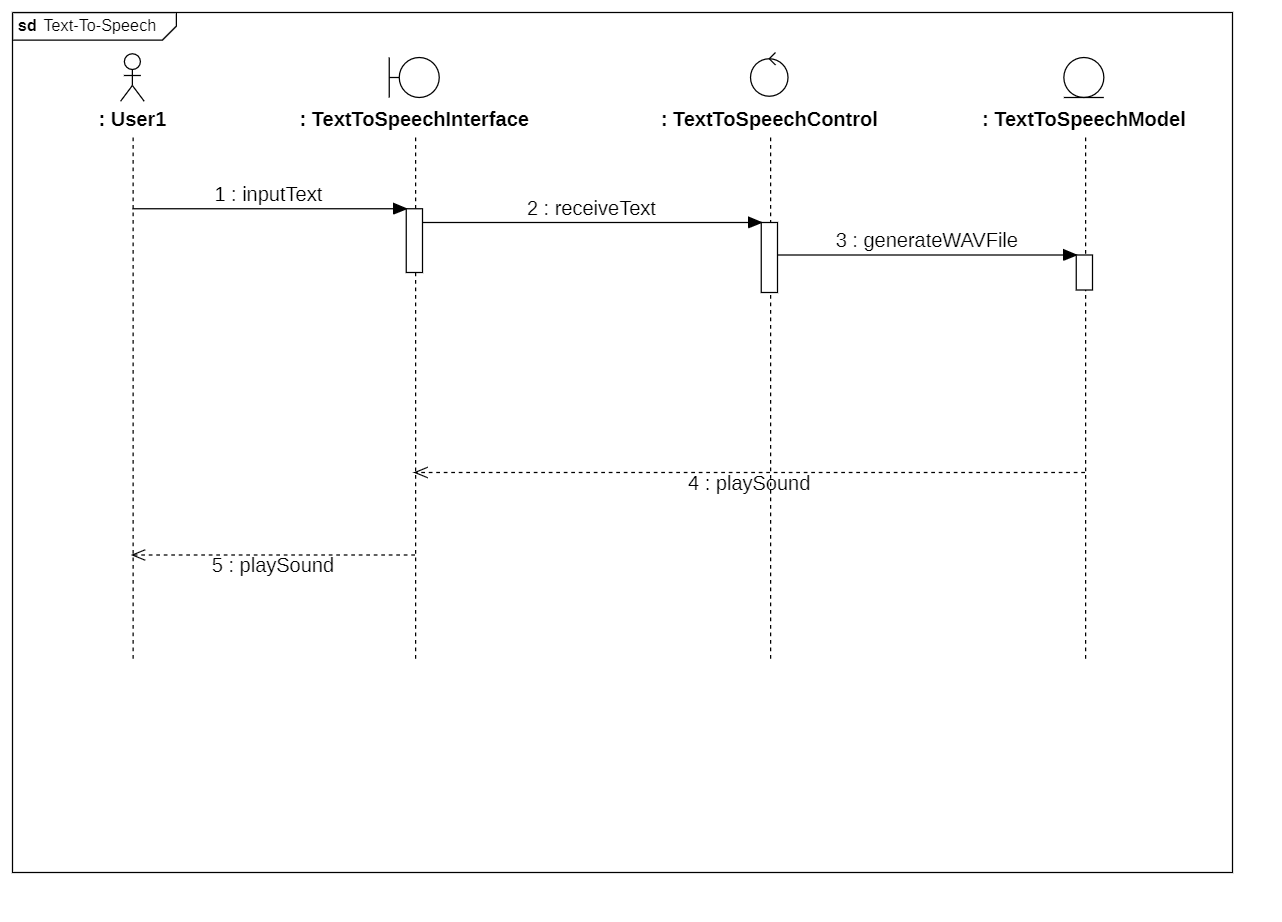
ຕາຕະລາງທີ 3.4 ຄໍາອະທິບາຍ Speech-To-Text

|  |  |
| --- | --- |
| Use Case Title: Speech-To-Text | Use Case ID: 2 |
| Primary Actor: User | |
| **Main Flow:** ໃນກໍລະນີທີ່ User ຕ້ອງການໃຊ້ Speech-To-Text, User ຕ້ອງປ້ອນ: ຟາຍສຽງ ເພື່ອໃຊ້ງານການແປງສຽງໄປເປັນຂໍ້ຄວາມ | |
| **Exceptional Flow 1:** ໃນກໍລະນີທີ່ User ປ້ອນຂໍ້ມູນບໍ່ຄົບຖ້ວນ, ລະບົບຈະແຈ້ງເຕືອນວ່າ “ກະລຸນາປ້ອນຂໍ້ມູນໃຫ້ຄົບຖ້ວນ”. | |

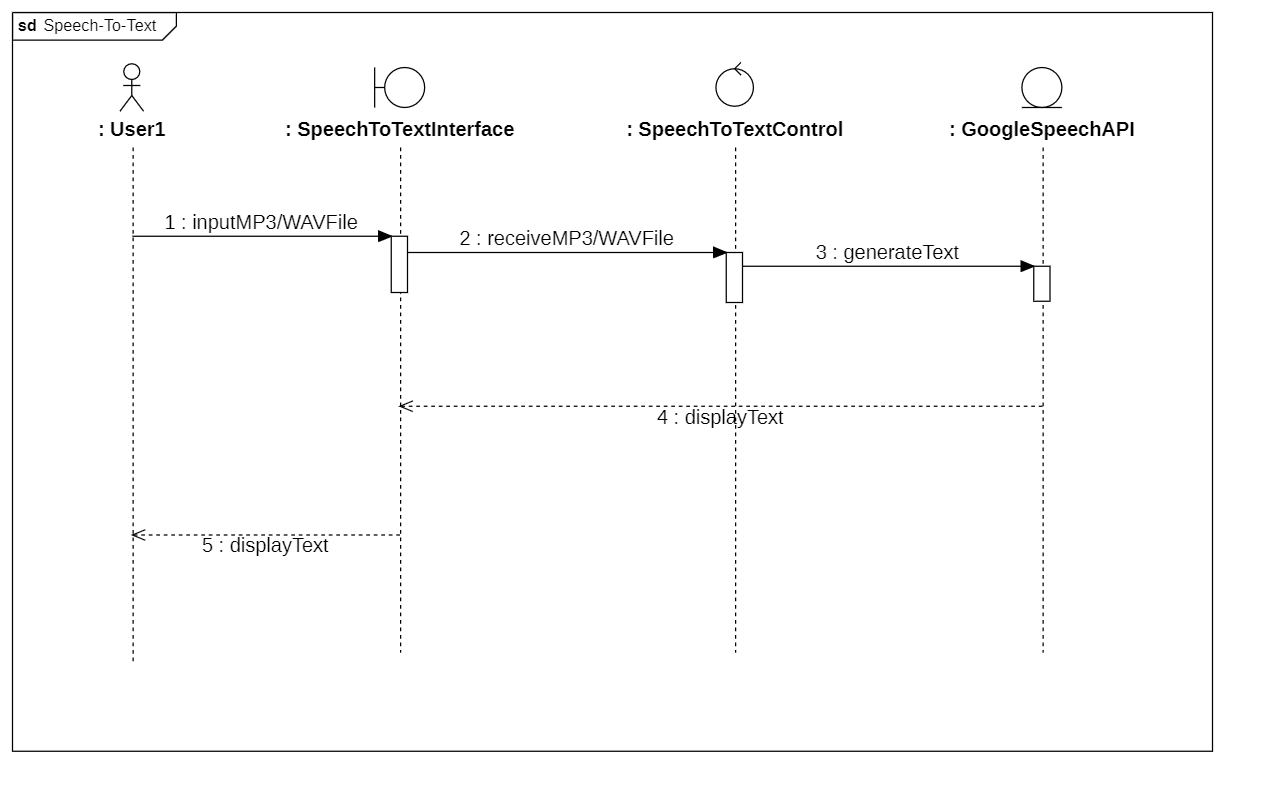
ຈາກ Main Flow ຂອງ Use Case Speech-To-Text ແມ່ນບໍ່ສາມາດຄົ້ນຫາ Class ແລະ Attribute ໄດ້ເນື່ອງຈາກວ່າບ່ໄດ້ຈັດເກັບຂໍ້ມູນທີ່ປ້ອນໄວ້.

### ແຜນວາດ Sequence Diagram

ຫຼັງຈາກໄດ້ຕາຕະລາງຄຳອະທິບາຍຂອງແຕ່ລະ Use Case ເຮົາສາມາດນໍາມາຂຽນເປັນແຜ່ນພາບ Sequence Diagram ດັ່ງລຸ່ມນີ້:



ແຜນວາດທີ 3.3 ແຜນວາດສະແດງ Sequence Diagram ຂອງ Text-To-Speech



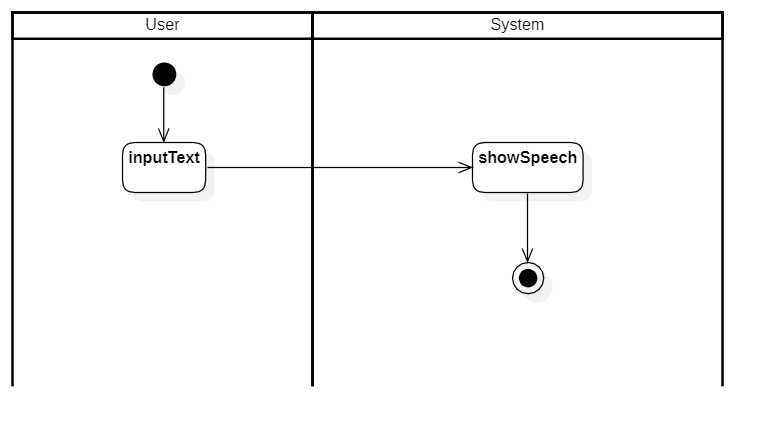
ແຜນວາດທີ 3.4 ແຜນວາດສະແດງ Sequence Diagram ຂອງ Speech-To-Text

### ແຜນວາດ Activity Diagram

#### ແຜນວາດ Activity ການແປຂໍ້ຄວາມຂຽນໄປເປັນຂໍ້ຄວາມສຽງ (Text-To-Speech)

ຂັ້ນຕອນກິດຈະກຳການແປຂໍ້ຄວາມຂຽນໄປເປັນຂໍ້ຄວາມສຽງມີດັ່ງລຸ່ມນີ້:

1. ເລີ່ມຕົ້ນການເຮັດວຽກ
2. ປ້ອນຂໍ້ຄວາມທີ່ຕ້ອງການແປຂໍ້ຄວາມຂຽນໄປເປັນຂໍ້ຄວາມສຽງ
3. ກວດສອບການປ້ອນຂໍ້ຄວາມວ່າງເປົ່າ ຫຼື ບໍ່ວ່າງເປົ່າ
4. ກໍລະນີຂໍ້ຄວາມວ່າງເປົ່າ ໃຫ້ປ້ອນຂໍ້ຄວາມໃໝ່
5. ກໍລະນີຂໍ້ຄວາມບໍ່ວ່າງເປົ່າ ໃຫ້ສະແດງຂໍ້ຄວາມສຽງອອກມາທາງເວັບໄຊ
6. ຈົບການເຮັດວຽກ

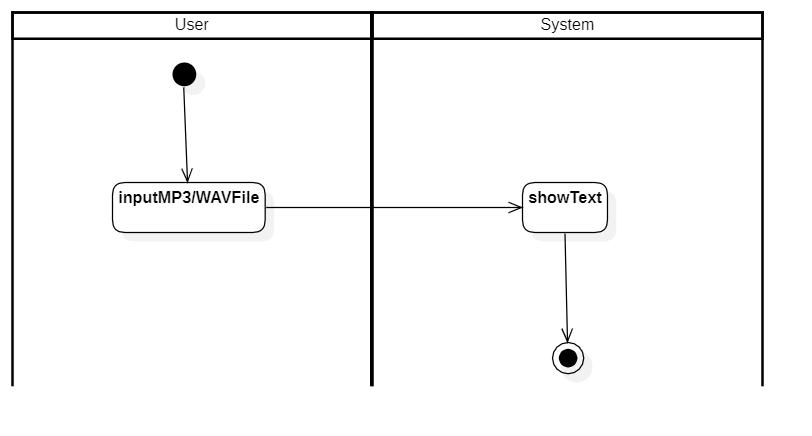


ແຜນວາດທີ 3.5 ແຜນວາດສະແດງ Activiti Diagram ຂອງ Text-To-Speech

#### ແຜນວາດ Activity ການແປຂໍ້ຄວາມສຽງໄປເປັນຂໍ້ຄວາມຂຽນ (Speech-To-Text)

ຂັ້ນຕອນກິດຈະກຳການແປຂໍ້ຄວາມສຽງໄປເປັນຂໍ້ຄວາມຂຽນມີດັ່ງລຸ່ມນີ້:

1. ເລີ່ມຕົ້ນການເຮັດວຽກ
2. ປ້ອນຟາຍສຽງທີ່ຕ້ອງການແປຂໍ້ຄວາມສຽງໄປເປັນຂໍ້ຄວາມຂຽນ
3. ກວດສອບການປ້ອນຟາຍສຽງວ່າງເປົ່າ ຫຼື ບໍ່ວ່າງເປົ່າ
4. ກໍລະນີຂໍ້ຄວາມວ່າງເປົ່າ ໃຫ້ປ້ອນຟາຍສຽງໃໝ່
5. ກໍລະນີຂໍ້ຄວາມບໍ່ວ່າງເປົ່າ ໃຫ້ສະແດງຂໍ້ຄວາມຂຽນອອກມາທາງເວັບໄຊ
6. ຈົບການເຮັດວຽກ



ແຜນວາດທີ 3.6 ແຜນວາດສະແດງ Activiti Diagram ຂອງ Speech-To-Text

## ການອອກແບບລະບົບ



### ການອອກແບບຮ່າງສະແດງຜົນ

ການພັດທະນາ Website ໃນບົດໂຄງການນີ້ພວກຂ້າພະເຈົ້າໄດ້ອອກແບບລວມເອົາທັງ TTS ແລະ STT ໃວ້ໃນໜ້າດຽວ (Page Website) ໂດຍການແບ່ງໜ້າວຽກອອກເປັນສ່ວນໆດັ່ງລຸ່ມນີ້:

#### ພາກສ່ວນການແປຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ

LTDeveloper

Project Title

1. ການແປຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ

Box ປ້ອນຂໍ້ຄວາມ

ປຸ່ມກົດຢືນຢັນ TTS

**ຜົນໄດ້ຮັບ:**

Audio

ປຸ່ມກົດບັນທຶກສຽງ

Logo

ຮູບທີ 3.8 ຮູບພາບສະແດງການອອກແບບພາກສ່ວນການແປຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ

#### ພາກສ່ວນການແປຂໍ້ຄວາມສຽງພາສາລາວໄປເປັນຂໍ້ຄວາມຂຽນ

2. ການແປຂໍ້ຄວາມສຽງພາສາລາວໄປເປັນຂໍ້ຄວາມຂຽນ

ກະລຸນາເລືອກຟາຍສຽງ:

Browse

ປຸ່ມກົດຢືນຢັນ STT

Box ສະແດງຂໍ້ຄວາມ

ອາຈານຜູ້ຊ່ວຍນຳພາ:

ອຈ.ປທ ສົມມິດ ທຸມມາລີ

ອາຈານຜູ້ນຳພາ:

ອຈ.ປອ ລັດສະໝີ ຈິດຕະວົງ

Copyright © LTDeveloper

Audio ຫຼັງເລືອກຟາຍສຽງ

ຮູບທີ 3.9 ຮູບພາບສະແດງການອອກແບບພາກສ່ວນການແປຂໍ້ຄວາມສຽງພາສາລາວໄປເປັນຂໍ້ຄວາມຂຽນ

### ການອອກແບບໂປຣແກຣມ

Input Text

Output Streaming Audio

End

Start

No

Yes

Text-To-Speech

ແຜນວາດທີ 3.7 ແຜນວາດ Flowchart ສະແດງເຖິງການເຮັດວຽກຂອງ Text-To-Speech

Start

Yes

No

Speech-To-Text

Input Speech

Output Text

Output Streaming Audio

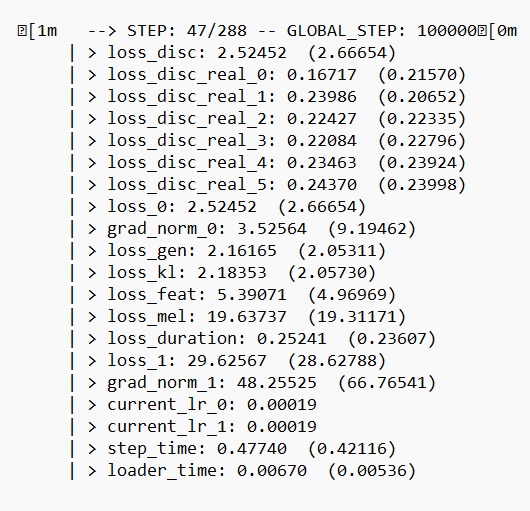
End

ແຜນວາດທີ 3.8 ແຜນວາດ Flowchart ສະແດງເຖິງການເຮັດວຽກຂອງ Speech-To-Text

# ຜົນຂອງການສຶກສາ ແລະ ການອະທິບາຍຜົນ



## ການລາຍງານຜົນຂອງການຄົ້ນຄວ້າ



ຮູບທີ 4.1 ຮູບພາບສະແດງຄ່າຕ່າງໆໃນການປະມວນຜົນແບບຈຳລອງໃນຮອບທີ 100,000

ຈາກຮູບທີ 4.1 ສະແດງໃຫ້ເຫັນວ່າ ຄ່າຕ່າງທີ່ວັດປະສິດທິພາບຂອງ model ເມື່ອຄົົບຈໍານວນທີ່ທົດສອບໄປ 100,000 ຮອບ ເຊິ່ງຈະຍົກມາສະເພາະຕົວທີ່ສໍາຄັນດັ່ງນີ້:

###### ຄ່າເວລາສະເລ່ຍໃນການໂຫຼດຂໍ້ມູນການໃນລະຫວ່າງການເຝິກສອນແບບຈຳລອງ loader\_time ທີ່ວັດຄ່າໄດ້ 0.00670 ວິນາທີ

###### ຄ່າໄລຍະເວລາການສູນເສຍການຄາດຄະເນແລະໄລຍະເວລາຄວາມຈິງພື້ນຖານຂອງ phoneme ໃນການເວົ້າທີ່ສ້າງຂຶ້ນໃນລະຫວ່າງການເຝິກສອນແບບຈຳລອງ loss\_duration ທີ່ວັດຄ່າໄດ້ 0.025241 ວິນາທີ

###### ຄ່າການສູນເສຍຄັ້ງທໍາອິດທີ່ໃຊ້ໃນລະຫວ່າງການເຝິກສອນຂອງ vocoder MelGAN loss\_01 ໄດ້ຄ່າຢູ່ທີ່ 29.62567

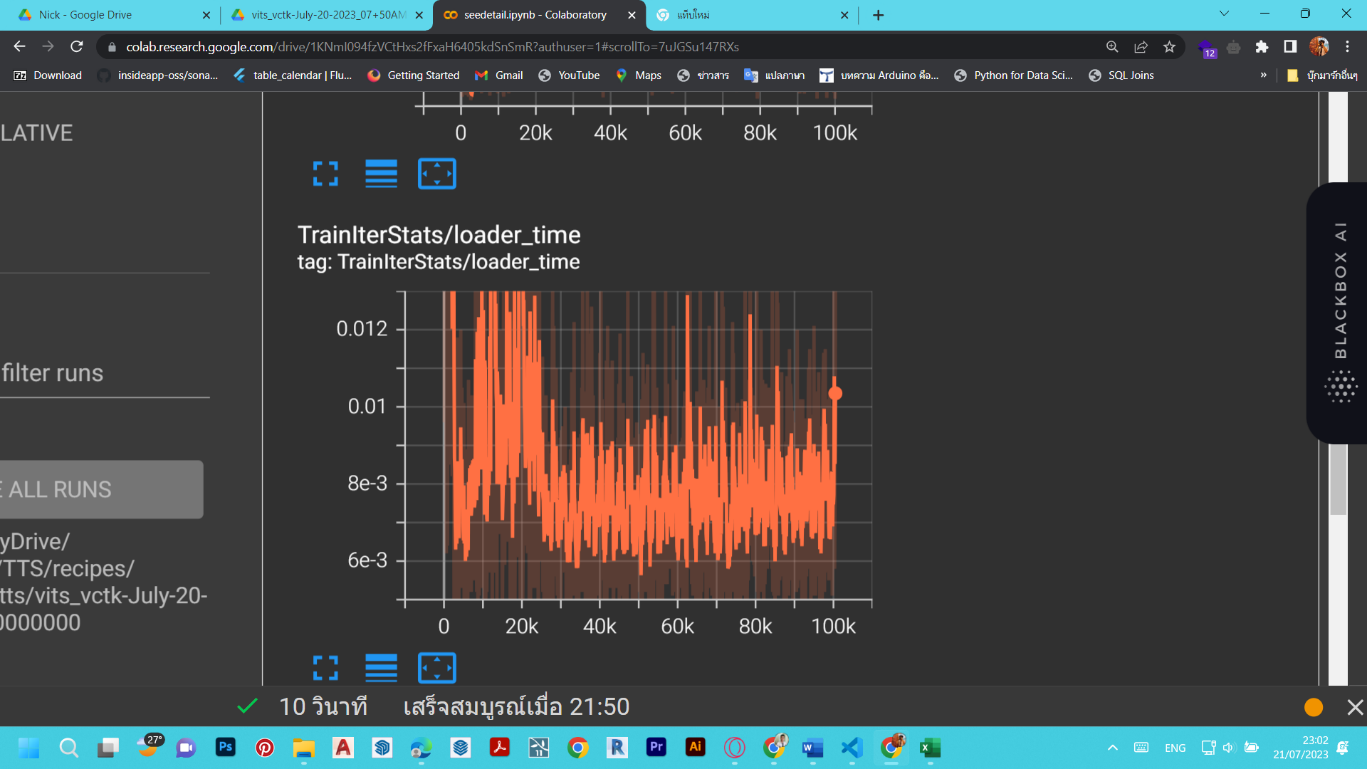
###### ຄ່າການສູນເສຍ mel-spectrogram ໃນລະຫວ່າງການເຝິກສອນແບບຈໍາລອງ ໄດ້ຄ່າຢູ່ທີ່ 19.63737



### ລາຍງານຜົນການສຶກສາຂັ້ນຕອນການເຝິກສອນແບບຈຳລອງ

ຜົນການສຶກສາຂັ້ນຕອນການເຝິກສອນແບບຈຳລອງການແປຂໍ້ຄວາມຂຽນໄປເປັນຂໍ້ຄວາມສຽງ ໄດ້ສະແດງຜົນໄດ້ຮັບອອກມາເປັນໃນຮູບ Graph ດັ່ງນີ້:

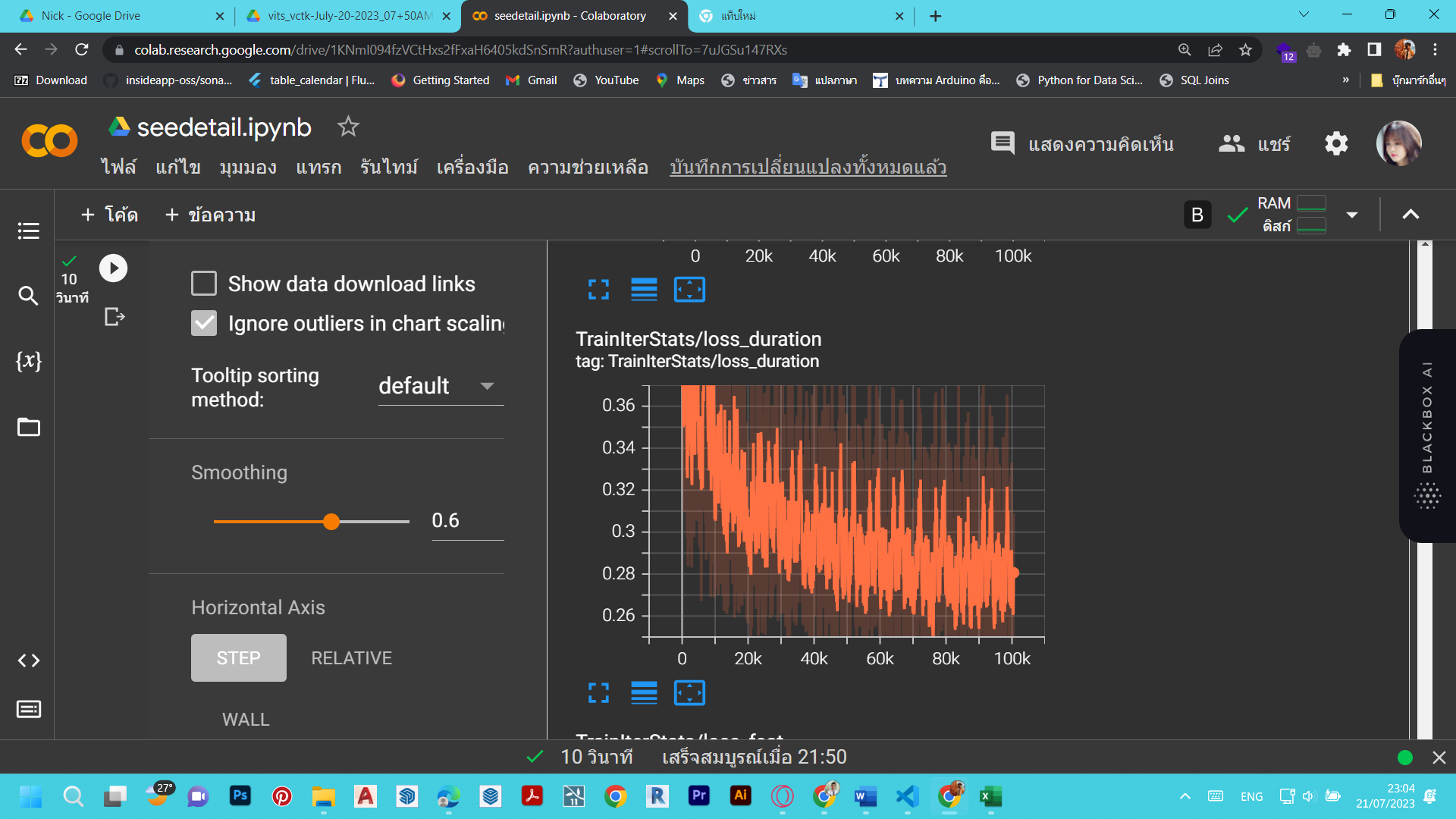
#### ເວລາສະເລ່ຍໃນການໂຫຼດຂໍ້ມູນການໃນລະຫວ່າງການເຝິກສອນແບບຈຳລອງ



ຮູບທີ 4.2 ຮູບພາບສະແດງເວລາສະເລ່ຍໃນການໂຫຼດຂໍ້ມູນການໃນລະຫວ່າງການເຝິກສອນແບບຈຳລອງ (TrainIterStats/loader\_time)

ຈາກຮູບທີ 4.2 ສະແດງໃຫ້ເຫັນວ່າ ເວລາສະເລ່ຍໃນການໂຫຼດຂໍ້ມູນການເຝິກສອນ ແຕ່ລະຊຸດມີທ່າອ່ຽງຫຼຸດລົງເລື້ອຍໆຈາກ 0.012 ວິນາທີໄດ້ກາຍມາເປັນ 0.009 ວິນາທີ ເມື່ອຄົົບຈຳນວນທີ່ທົດສອບໄປ 100,000 ຮອບ, ແຕ່ຄ່າທີ່ຕໍ່າທີ່ສຸດ ຈະຢູ່ທີ່ຮອບທີ່ 50,721 ທີ່ວັດຄ່າໄດ້ຢູ່ທີ່ 0.005 ວິນາທີ.

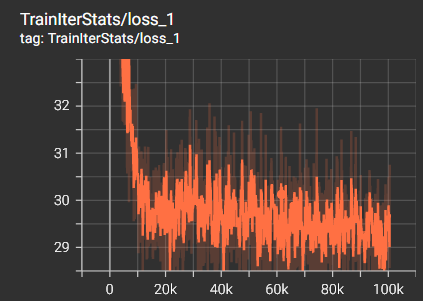
#### ໄລຍະເວລາການສູນເສຍການຄາດຄະເນແລະໄລຍະເວລາຄວາມຈິງພື້ນຖານຂອງ phoneme ໃນການເວົ້າທີ່ສ້າງຂຶ້ນໃນລະຫວ່າງການເຝິກສອນແບບຈຳລອງ



ຮູບທີ 4.3 ຮູບພາບສະແດງໄລຍະເວລາການສູນເສຍການຄາດຄະເນແລະໄລຍະເວລາຄວາມຈິງພື້ນຖານຂອງ phoneme ໃນການເວົ້າທີ່ສ້າງຂຶ້ນໃນລະຫວ່າງການເຝິກສອນແບບຈຳລອງ (TrainIterStats/loss\_duration)

ຈາກຮູບທີ 4.3 ສະແດງໃຫ້ເຫັນວ່າ ໄລຍະເວລາການສູນເສຍການຄາດຄະເນແລະໄລຍະເວລາຄວາມຈິງພື້ນຖານຂອງ phoneme ໃນການເວົ້າທີ່ສ້າງຂຶ້ນ ມີທ່າອ່ຽງຫຼຸດລົງເຊັ່ນດຽວກັນຈາກ 0.424 ວິນາທີ ກາຍມາເປັນ 0.2806 ວິນາທີ ເມື່ອຄົົບຈໍານວນທີ່ທົດສອບໄປ 100,000 ຮອບ, ແຕ່ຄ່າທີ່ຕໍ່າທີ່ສຸດ ຈະຢູ່ທີ່ຮອບທີ່ 76,100 ທີ່ວັດຄ່າໄດ້ຢູ່ທີ່ 0.246 ວິນາທີ.

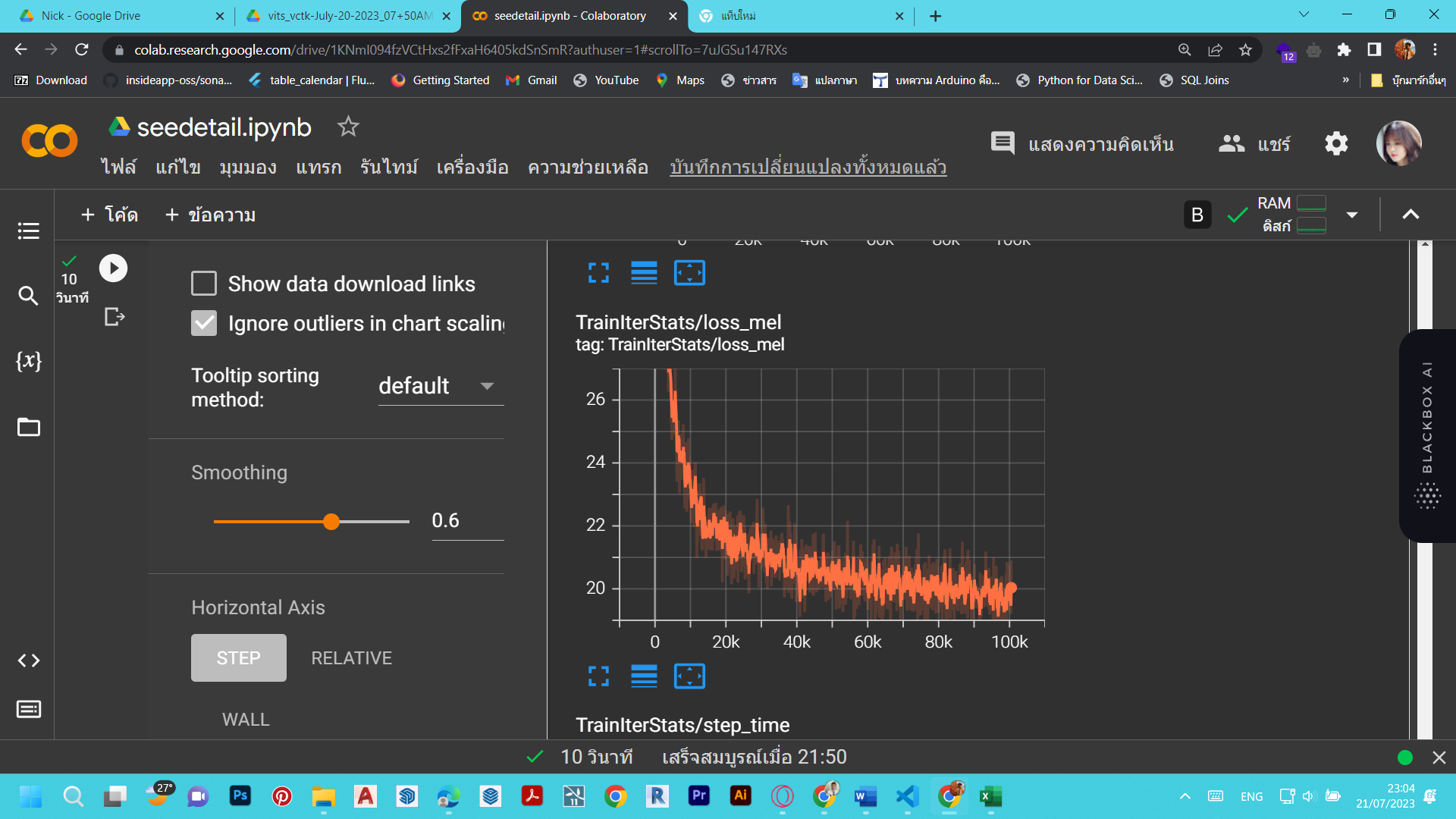
#### ການສູນເສຍຄັ້ງທໍາອິດທີ່ໃຊ້ໃນລະຫວ່າງການເຝິກສອນຂອງ vocoder MelGAN



ຮູບທີ 4.4 ຮູບພາບສະແດງການສູນເສຍຄັ້ງທໍາອິດທີ່ໃຊ້ໃນລະຫວ່າງການເຝິກສອນຂອງ vocoder MelGAN (TrainIterStats/loss\_01)

ຈາກຮູບທີ 4.4 ສະແດງໃຫ້ເຫັນວ່າ ການສູນເສຍຄັ້ງທໍາອິດທີ່ໃຊ້ໃນລະຫວ່າງການຝຶກອົບຮົມຂອງ vocoder MelGAN ມີທ່າອ່ຽງຫຼຸດລົງ ເຊັ່ນດຽວກັນຈາກ 52.58 ຫຼຸດລົງກາຍມາເປັນ 29.71 ເມື່ອຄົົບຈໍານວນທີ່ທົດສອບໄປ 100,000 ຮອບ, ແຕ່ຄ່າທີ່ຕໍ່າທີ່ສຸດ ຈະຢູ່ທີ່ຮອບທີ່ 40,900 ທີ່ວັດຄ່າໄດ້ຢູ່ທີ່ 28.28.

#### ການສູນເສຍ mel-spectrogram ໃນລະຫວ່າງການເຝິກສອນແບບຈຳລອງ



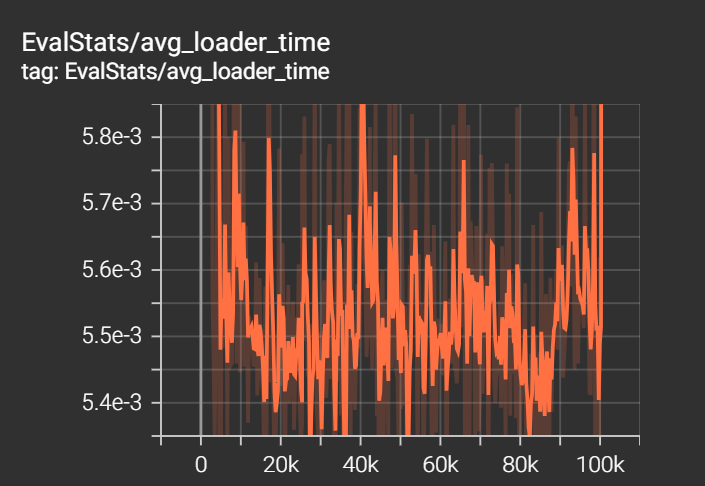
ຮູບທີ 4.5 ຮູບພາບສະແດງການສູນເສຍ mel-spectrogram ໃນລະຫວ່າງການເຝິກສອນແບບຈຳລອງ (TrainIterStats/loss\_mel)

ຈາກຮູບທີ 4.5 ສະແດງໃຫ້ເຫັນວ່າ ການສູນເສຍ mel-spectrogram ມີທ່າອ່ຽງຫຼຸດລົງ ເຊັ່ນດຽວກັນຈາກ 50 ຫຼຸດລົງກາຍມາເປັນ 19.7 ເມື່ອຄົົບຈໍານວນທີ່ທົດສອບໄປ 100,000 ຮອບ, ແຕ່ຄ່າທີ່ຕໍ່າທີ່ສຸດ ຈະຢູ່ທີ່ຮອບທີ່ 91,600 ທີ່ວັດຄ່າໄດ້ຢູ່ທີ່ 19.2.

### ລາຍງານຜົນການສຶກສາຂັ້ນຕອນການທົດລອງແບບຈຳລອງ

ຜົນການສຶກສາຂັ້ນຕອນການທົດລອງແບບຈຳລອງການແປຂໍ້ຄວາມຂຽນໄປເປັນຂໍ້ຄວາມສຽງ ໄດ້ສະແດງຜົນໄດ້ຮັບອອກມາເປັນໃນຮູບ Graph ດັ່ງນີ້:

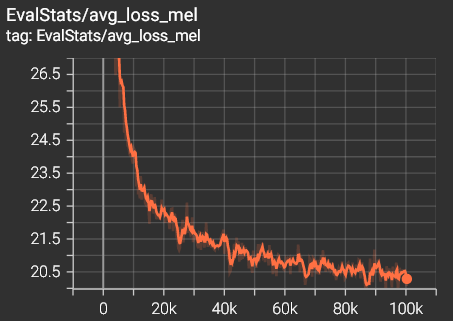
#### ເວລາສະເລ່ຍໃນການໂຫຼດຂໍ້ມູນໃນລະຫວ່າງການທົດລອງ



ຮູບທີ 4.6 ຮູບພາບສະແດງເວລາສະເລ່ຍໃນການໂຫຼດຂໍ້ມູນໃນລະຫວ່າງການທົດລອງ (EvalStats/avg\_loader\_time)

ຈາກຮູບທີ 4.6 ສະແດງໃຫ້ເຫັນວ່າ ເວລາສະເລ່ຍໃນການໂຫຼດຂໍ້ມູນໃນລະຫວ່າງການທົດລອງ ແຕ່ລະຊຸດມີທ່າອ່ຽງຫຼຸດລົງເລື້ອຍໆຈາກ 0.008 ວິນາທີໄດ້ກາຍມາເປັນ 0.0054 ວິນາທີ ເມື່ອຄົົບຈຳນວນທີ່ທົດສອບໄປ 100,000 ຮອບ, ແຕ່ຄ່າທີ່ຕໍ່າທີ່ສຸດ ຈະຢູ່ທີ່ຮອບທີ່ 27,370 ທີ່ວັດຄ່າໄດ້ຢູ່ທີ່ 0.0053 ວິນາທີ.

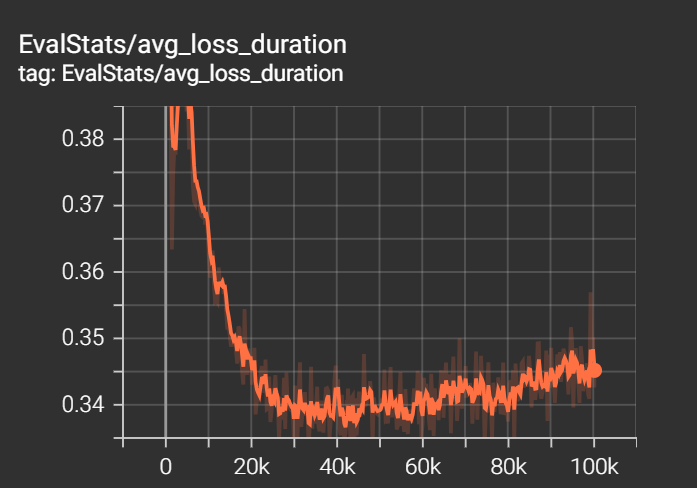
#### ການສູນເສຍ mel-spectrogram ໃນລະຫວ່າງການທົດລອງ



ຮູບທີ 4.7 ຮູບພາບສະແດງການສູນເສຍ mel-spectrogram (EvalStats/avg\_loss\_mel)

ຈາກຮູບທີ 4.7 ສະແດງໃຫ້ເຫັນວ່າ ການສູນເສຍ mel-spectrogram ໃນລະຫວ່າງການທົດລອງ ມີທ່າອ່ຽງຫຼຸດລົງ ເຊັ່ນດຽວກັນຈາກ 37.5 ຫຼຸດລົງກາຍມາເປັນ 20.29 ເມື່ອຄົົບຈຳນວນທີ່ທົດສອບໄປ 100,000 ຮອບ, ແຕ່ຄ່າທີ່ຕໍ່າທີ່ສຸດ ຈະຢູ່ທີ່ຮອບທີ່ 87,000 ທີ່ວັດຄ່າໄດ້ຢູ່ທີ່ 20.11.

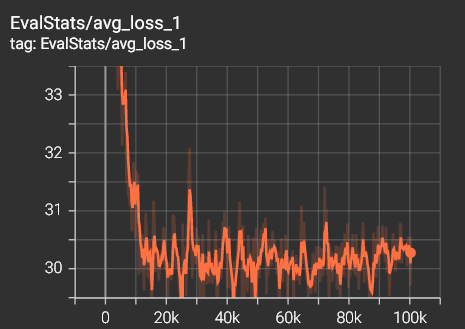
#### ໄລຍະເວລາການສູນເສຍການຄາດຄະເນແລະໄລຍະເວລາຄວາມຈິງພື້ນຖານຂອງ phoneme ໃນການເວົ້າທີ່ສ້າງຂຶ້ນໃນລະຫວ່າງການທົດລອງ



ຮູບທີ 4.8 ຮູບພາບສະແດງໄລຍະເວລາການສູນເສຍການຄາດຄະເນ ແລະ ໄລຍະເວລາຄວາມຈິງພື້ນຖານຂອງ phoneme ໃນການເວົ້າທີ່ສ້າງຂຶ້ນ (EvalStats/ avg\_loss\_duration)

ຈາກຮູບທີ 4.8 ສະແດງໃຫ້ເຫັນວ່າໄລຍະເວລາການສູນເສຍການຄາດຄະເນແລະໄລຍະເວລາຄວາມຈິງພື້ນຖານຂອງ phoneme ໃນການເວົ້າທີ່ສ້າງຂຶ້ນໃນລະຫວ່າງການທົດລອງ ມີທ່າອ່ຽງຫຼຸດລົງເຊັ່ນດຽວກັນຈາກ 0.412 ວິນາທີ ກາຍມາເປັນ 0.345 ວິນາທີ ເມື່ອຄົົບຈຳນວນທີ່ທົດສອບໄປ 100,000 ຮອບ, ແຕ່ຄ່າທີ່ຕໍ່າທີ່ສຸດ ຈະຢູ່ທີ່ຮອບທີ່ 42,050 ທີ່ວັດຄ່າໄດ້ຢູ່ທີ່ 0.336 ວິນາທີ.

#### ການສູນເສຍ mel-spectrogram ໃນລະຫວ່າງການທົດລອງ



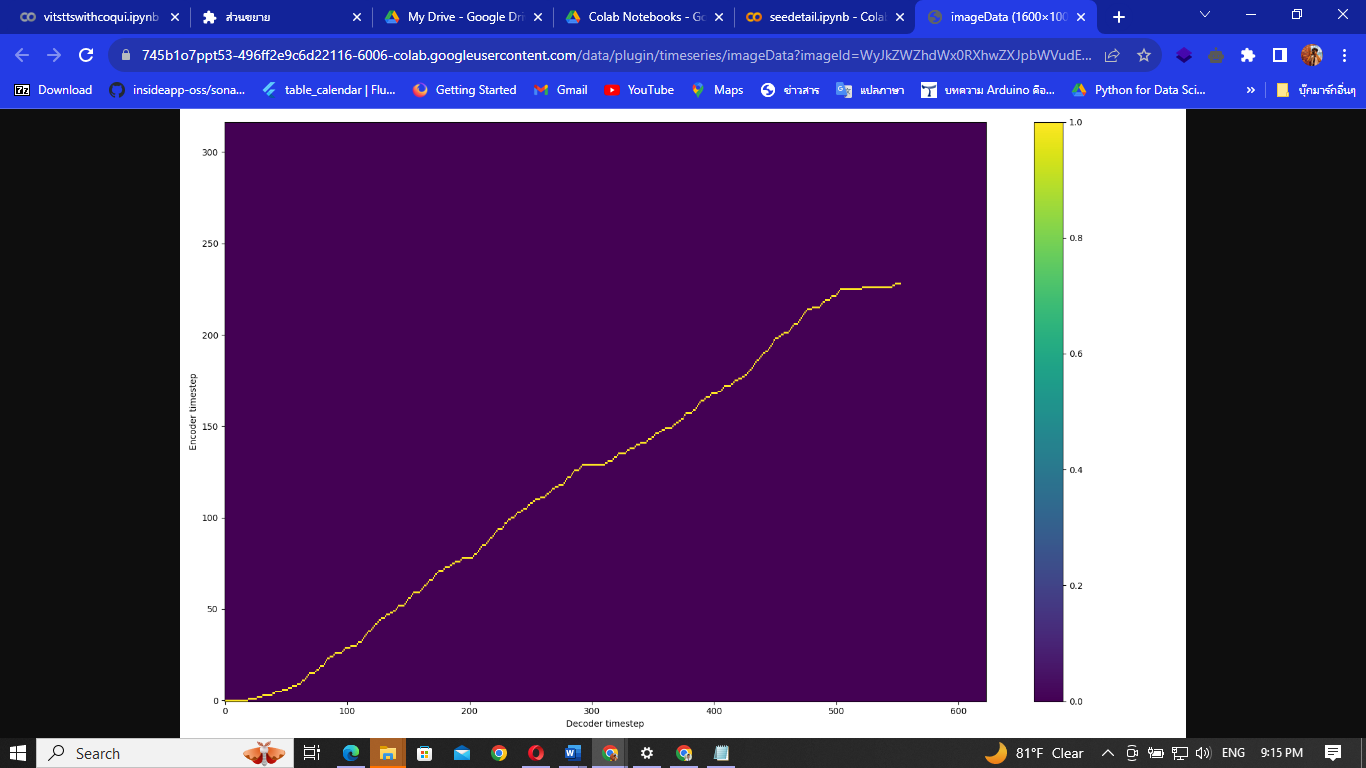
ຮູບທີ 4.9 ຮູບພາບສະແດງການສູນເສຍ mel-spectrogram ໃນລະຫວ່າງການທົດລອງ (EvalStats/ avg\_loss\_1)

ຈາກຮູບທີ 4.9 ສະແດງໃຫ້ເຫັນວ່າ ການສູນເສຍ mel-spectrogram ໃນລະຫວ່າງການທົດລອງ ມີທ່າອ່ຽງຫຼຸດລົງ ເຊັ່ນດຽວກັນຈາກ 43.97 ຫຼຸດລົງກາຍມາເປັນ 30.27 ເມື່ອຄົົບຈຳນວນທີ່ທົດສອບໄປ 100,000 ຮອບ, ແຕ່ຄ່າທີ່ຕໍ່າທີ່ສຸດ ຈະຢູ່ທີ່ຮອບທີ່ 25.35 ທີ່ວັດຄ່າໄດ້ຢູ່ທີ່ 29.28.

### ລາຍງານຜົນການສຶກສາການປະເມີນແບບຈຳລອງ

ຜົນການສຶກສາຂັ້ນຕອນການປະເມີນແບບຈຳລອງການແປຂໍ້ຄວາມຂຽນໄປເປັນຂໍ້ຄວາມສຽງ ໄດ້ສະແດງຜົນໄດ້ຮັບອອກມາເປັນໃນຮູບ Graph ດັ່ງນີ້:

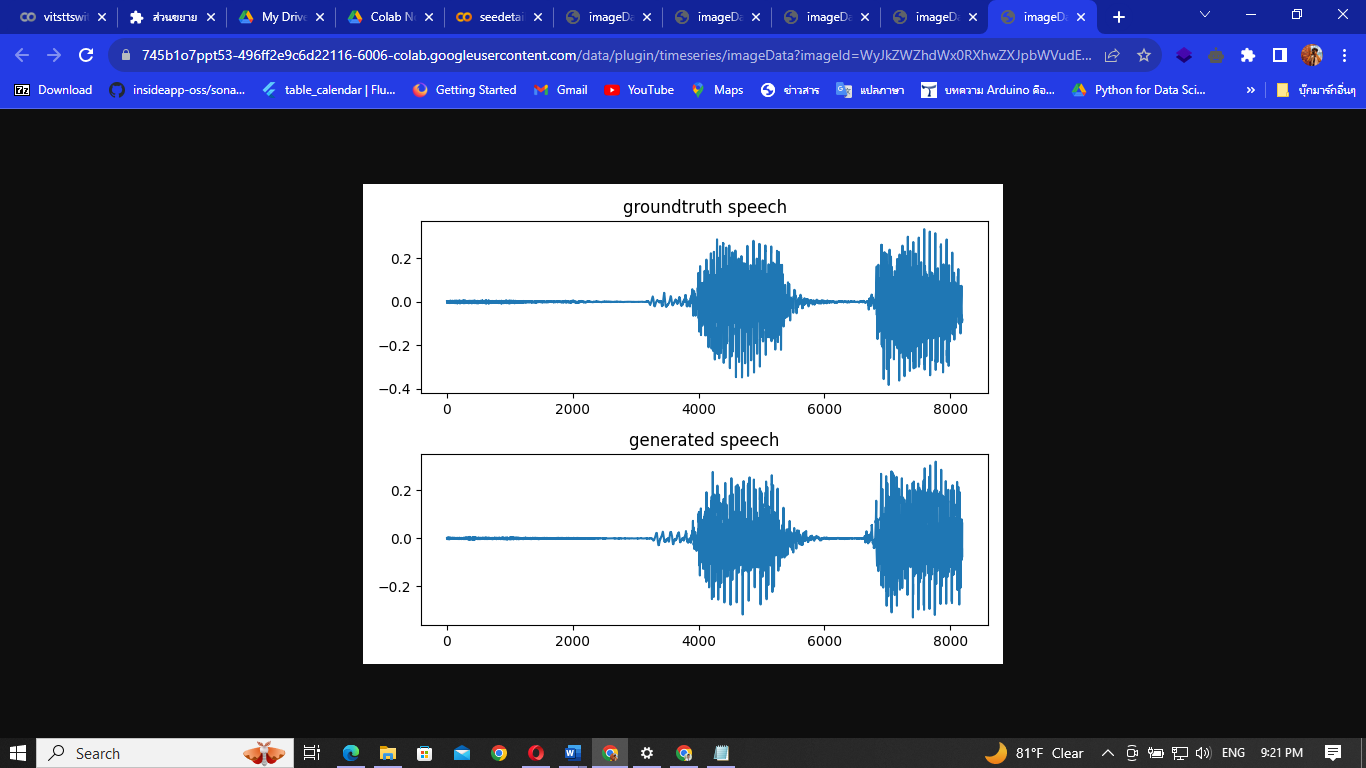
#### ຕົວຊີ້ວັດທີ່ໃຊ້ໃນແບບຈຳລອງ ເພື່ອປະເມີນຄຸນນະພາບຂອງການຈັດຮຽງລະຫວ່າງຂໍ້ຄວາມທີ່ປ້ອນເຂົ້າ ແລະ ຮູບແບບຄື້ນສຽງທີ່ສ້າງຂຶ້ນ



ຮູບທີ 4.10 ຮູບພາບສະແດງຕົວຊີ້ວັດທີ່ໃຊ້ໃນແບບຈຳລອງ ເພື່ອປະເມີນຄຸນນະພາບຂອງການຈັດຮຽງລະຫວ່າງຂໍ້ຄວາມທີ່ປ້ອນເຂົ້າ ແລະ ຮູບແບບຄື້ນສຽງທີ່ສ້າງຂຶ້ນ

ຈາກຮູບທີ 4.10 ສະແດງໃຫ້ເຫັນວ່າ ຕົວຊີ້ວັດທີ່ໃຊ້ໃນແບບຈຳລອງ ເພື່ອປະເມີນຄຸນນະພາບຂອງການຈັດຮຽງລະຫວ່າງຂໍ້ຄວາມທີ່ປ້ອນເຂົ້າ ແລະ ຮູບແບບຄື້ນສຽງທີ່ສ້າງຂຶ້ນ ມີຄ່າເພີ່ມຂຶ້ນ ຢ່າງເປັນໄລຍະສຳຄັນຈາກ 0 ໄປເຖິງ 200 ສໍາລັບຄ່າ Encoder Timesleep (ຕົວເຂົ້າລະຫັດ ການປ່ຽນຂໍ້ຄວາມທີ່ປ້ອນເຂົ້າໄປເປັນຄື້ນສຽງເວົ້າ.) ແລະ 0 ຫາ 557 ສໍາລັບຄ່າ Decode Timesleep (ຕົວຖອດລະຫັດ ການປ່ຽນຂໍ້ຄວາມທີ່ປ້ອນເຂົ້າໄປເປັນຄື້ນສຽງເວົ້າ).

#### ຕົວຊີ້ວັດທີ່ໃຊ້ ເພື່ອປຽບທຽບຄຸນນະພາບຂອງສຽງເວົ້າທີ່ສັງເຄາະດ້ວຍສັນຍານສຽງເວົ້າຕົ້ນສະບັບ



ຮູບທີ 4.11 ຮູບພາບສະແດງຕົວຊີ້ວັດທີ່ໃຊ້ ເພື່ອປຽບທຽບຄຸນນະພາບຂອງສຽງເວົ້າທີ່ສັງເຄາະດ້ວຍສັນຍານສຽງເວົ້າຕົ້ນສະບັບ

ຈາກຮູບທີ 4.11 ສະແດງໃຫ້ເຫັນວ່າຕົວຊີ້ວັດທີ່ໃຊ້ ເພື່ອປຽບທຽບຄຸນນະພາບຂອງສຽງເວົ້າທີ່ສັງເຄາະດ້ວຍສັນຍານສຽງເວົ້າຕົ້ນສະບັບ ເຊິ່ງເຫັນວ່າຄ່າ groundtext speech (ສັນຍານການເວົ້າອ້າງອີງທີ່ຖືກນໍາໃຊ້ເພື່ອປະເມີນຄຸນນະພາບຂອງສຽງເວົ້າທີ່ສັງເຄາະ) ຈະເຫັນໄດ້ວ່າຄ່າຈະຢຸດຕັ້ງແຕ່ 8,000 ຂຶ້ນໄປໝາຍຄວາມວ່າຄຸນນະພາບສຽງທີ່ສັງເຄາະເລີ່ມມີປະສິດທິພາບຕັ້ງແຕ່ 8,000 ຮອບແລ້ວພຽງແຕ່ຈະເຝິກຕໍ່ໃຫ້ມີປະສິດທິພາບສຽງດີຂຶ້ນກວ່ານີ້

### ລາຍງານຜົນການຂຽນເວັບໄຊ ແລະ ວິທີການນຳໃຊ້

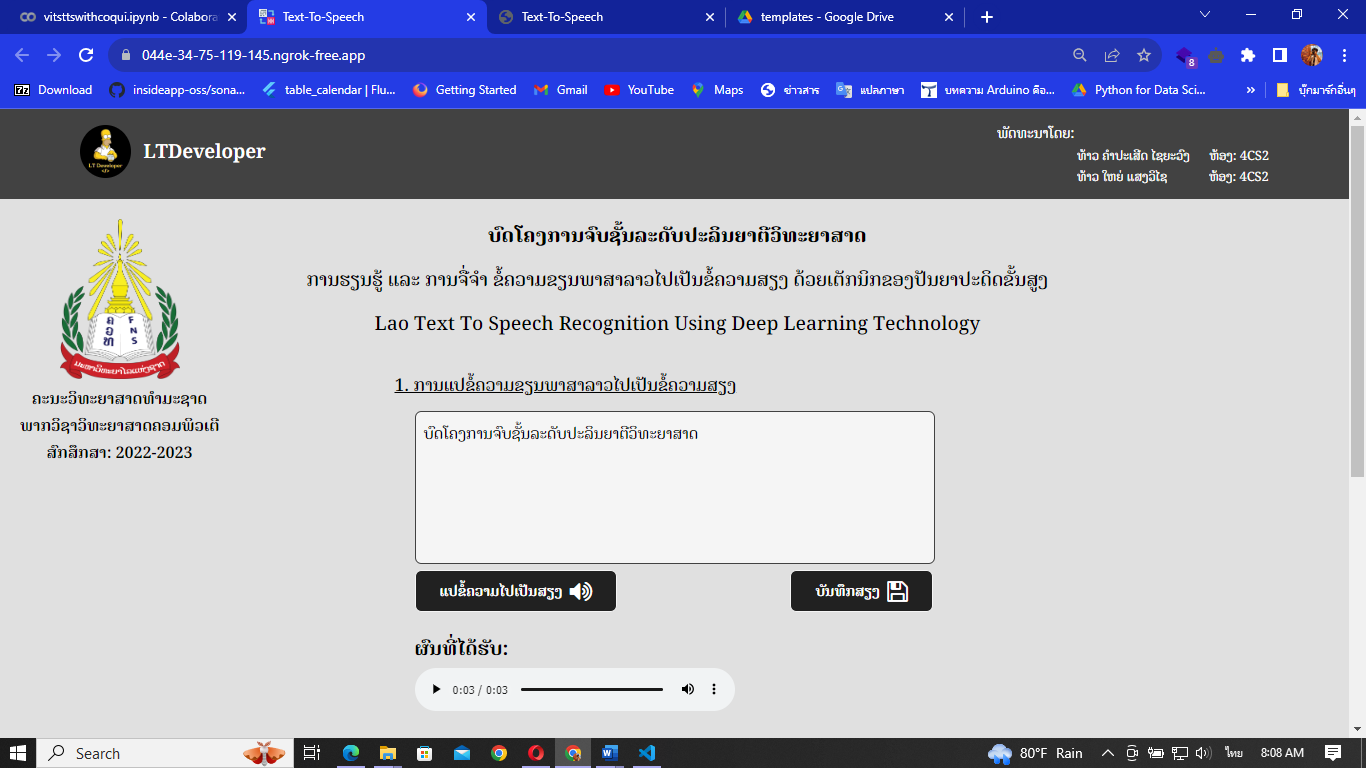
#### ລາຍງານຜົນການຂຽນເວັບໄຊ

ການພັດທະນາເວັບໄຊ ແມ່ນໄດ້ແບ່ງອອກເປັນ 2 ພາກສ່ວນຄື: ພາກສ່ວນ Back End ແມ່ນໄດ້ນຳເອົາ Flask Framework ໃນການພັດທະນາໂດຍນໍາໃຊ້ພາສາ Python ແລະ ພາກສ່ວນ Front End ແມ່ນໄດ້ນໍາໃຊ້ພາສາ HTML, CSS ແລະ JavaScript. ສໍາລັບເຄື່ອງມືທີ່ໃຊ້ຂຽນເວັບໄຊແມ່ນ Visual Studio Code.

#### ຂັ້ນຕອນວິທີການນຳໃຊ້ເວັບໄຊ

ຂັ້ນຕອນວິທີການນຳໃຊ້ເວັບໄຊ ແມ່ນຈະໄດ້ແນະນຳການນຳໃຊ້ຢູ່ 2 ພາກສ່ວນຄື: ພາກສ່ວນການແປຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ ແລະ ພາກສ່ວນການແປຂໍ້ຄວາມສຽງພາສາລາວໄປເປັນຂໍ້ຄວາມຂຽນ. ດັ່ງລຸ່ມນີ້:

##### ກ. ພາກສ່ວນການແປຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ



**1**

**4**

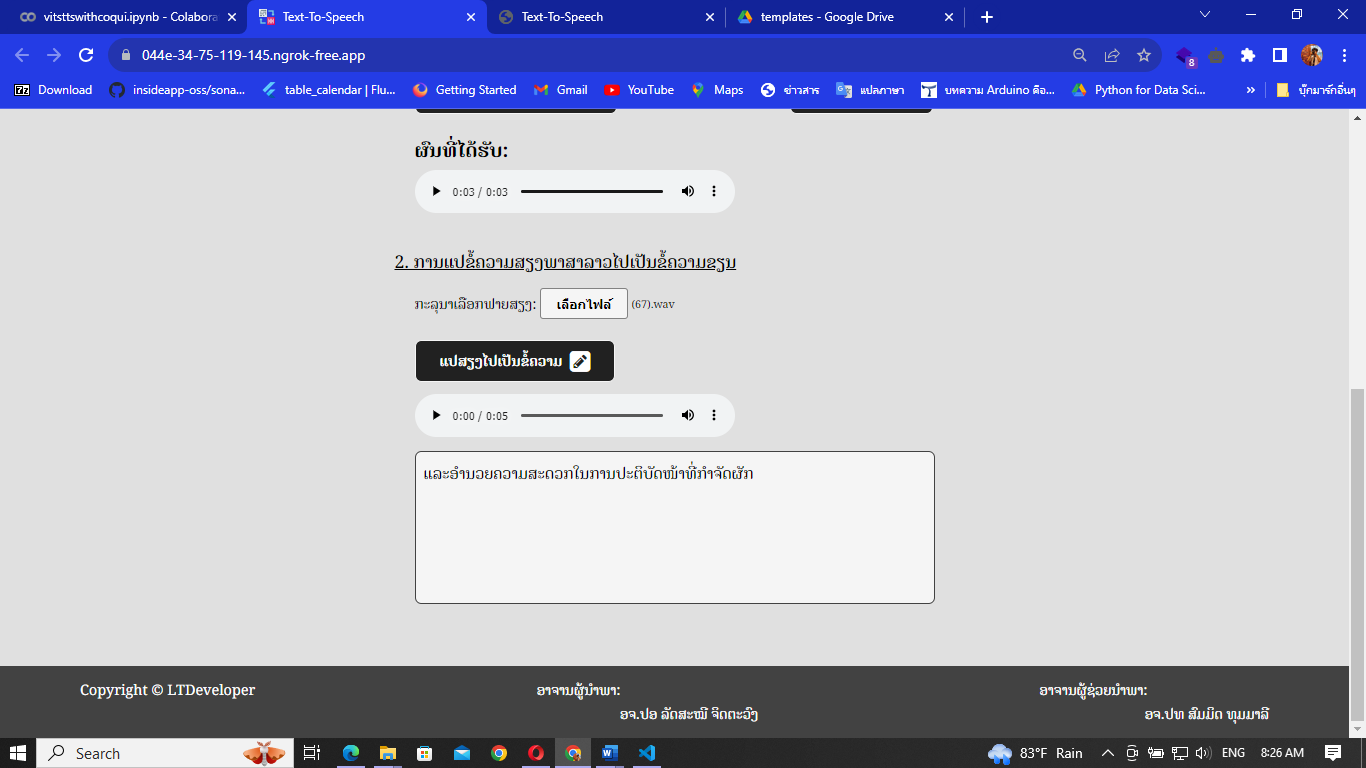
**2**

**3**

ຮູບທີ 4.12 ຮູບພາບສະແດງຂັ້ນຕອນວິທີການນໍາໃຊ້ເວັບໄຊໃນພາກສ່ວນການແປຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ

* 1. ປ້ອນຂໍ້ຄວາມພາສາລາວທີ່ທ່ານຕ້ອງການໃຫ້ເວັບໄຊແປໄປເປັນຂໍ້ຄວາມສຽງໃນຊ່ອງປ້ອນຂໍ້ຄວາມ
  2. ປຸ່ມກົດການແປຂໍ້ຄວາມໄປເປັນສຽງ ເພື່ອໃຫ້ເວັບໄຊເຮັດການແປຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນສຽງ
  3. ເຄື່ອງຫຼິ້ນສຽງ (Audio) ຈະສະແດງຂຶ້ນມາໃຫ້ເຫັນເມື່ອທ່ານກົດປຸ່ມການແປຂໍ້ຄວາມໄປເປັນສຽງແລ້ວ ທ່ານສາມາດກົດຟັງສຽງໃນເວັບໄຊໄດ້ທັນທີ
  4. ປຸ່ມບັນທຶກຟາຍສຽງຂໍ້ຄວາມພາສາລາວທີ່ທ່ານໄດ້ປ້ອນກ່ອນການກົດປຸ້ມການແປຂໍ້ຄວາມໄປເປັນສຽງຄັ້ງລ່າສຸດ

ຂ. ພາກສ່ວນການແປຂໍ້ຄວາມສຽງພາສາລາວໄປເປັນຂໍ້ຄວາມຂຽນ



**1**

**2**

**3**

**4**

ຮູບທີ 4.13 ຮູບພາບສະແດງຂັ້ນຕອນວິທີການນໍາໃຊ້ເວັບໄຊໃນພາກສ່ວນການແປຂໍ້ຄວາມສຽງພາສາລາວໄປເປັນຂໍ້ຄວາມຂຽນ

* 1. ກົດປຸ່ມເພື່ອເລືກຟາຍສຽງທີ່ທ່ານຕ້ອງການຈາກເຄື່ອງຄອມພິວເຕີຂອງທ່ານ
  2. ກົດປຸ່ມແປສຽງໄປເປັນຂໍ້ຄວາມ ເພື່ອໃຫ້ເວັບໄຊເຮັດການແປສຽງທີ່ທ່ານເລືອກຟາຍສຽງລ່າສຸດ
  3. ເຄື່ອງຫຼິ້ນສຽງ (Audio) ຈະສະແດງຂຶ້ນນມາໃຫ້ເຫັນ ເມື່ອທ່ານກົດປຸ່ມແປສຽງໄປເປັນຂໍ້ຄວາມແລ້ວ ທ່ານສາມາດກົດຟັງສຽງໃນເວັບໄຊໄດ້ທັນທີ
  4. ຫຼັງຈາກທີ່ທ່ານກົດປຸ່ມແປສຽງໄປເປັນຂໍ້ຄວາມແລ້ວ ຊ່ອງສະແດງຂໍ້ຄວາມຂໍ້ຈະສະແດງຄວາມຂຽນພາສາລາວອອກມາໃຫ້ເຫັນໃນຊ່ອງນີ້

## ການອະທິບາຍຜົນຂອງການຄົ້ນຄວ້າ

ຜົນການສຶກສາຄົ້ນຄ້ວາໃນຄັ້ງນີ້ເຫັນວ່າ ຜົນໄດ້ຮັບຂອງການປະມວນຜົນໃນການຮຽນຮູ້ ແລະ ຈື່ຈຳ ຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ ຂອງສະຖາປັດຕະຍະກຳຂອງ DNN ແລະ RNN ກໍຄືແບບຈຳລອງ YourTTS ສາມາດຈໍາແນກໄດ້ຄ່າຄວາມສູນເສຍໃນລະຫວ່າງການເຝິກສອນໄດ້ຄິດເປັນຄ່າຕ່າງໆຄື: ເວລາສະເລ່ຍໃນການໂຫຼດຂໍ້ມູນ (avg\_loader\_time) ມີຄ່າ 0.0054 ວິນາທີ, ການສູນເສຍ mel-spectrogram (avg\_loss\_mel) ມີຄ່າ 20.29, ເວລາການສູນເສຍການຄາດຄະເນ ແລະ ໄລຍະເວລາຄວາມຈິງພື້ນຖານຂອງ phoneme ໃນການເວົ້າທີ່ສ້າງຂຶ້ນ (avg\_loss\_duration) ມີຄ່າ 0.34 ວິນາທີ ແລະ ການສູນເສຍ mel-spectrogram (avg\_loss\_1) ມີຄ່າ 30.27 ໃນການເຝິກສອນໃນຮອບທີ່ 100,000.

ການສ້າງເວັບໄຊແມ່ນໄດ້ນຳໃຊ້ Flask Framework ໃນການເອີ້ນໃຊ້ແບບຈໍາລອງຂອງ YourTTS ທີ່ໄດ້ບັນທຶກອອກມາເປັນຟາຍ checkpoint\_100000.pth, speakers.pth ແລະ config.json, ພາສາໂປຣແກຣມທີ່ໄດ້ນໍາໃຊ້ໃນການສ້າງເວັບໄຊຫຼັກໆແມ່ນ ພາສາ Python ໃຊ້ເປັນໂຕຄວບຄຸມການເຮັດວຽກ ແລະ ປະມວນຜົນ. ສ່ວນເຄື່ອງມືທີ່ໃຊ້ໃນການສ້າງເວັບໄຊຫຼັກໆແມ່ນ Google Colab ໃຊ້ເປັນເຄື່ອງມືຊ່ວຍໃນການປະມວນຜົນ, ເຝິກສອນ ແລະ ທົດສອບແບບຈຳລອງ.

# ສະຫຼຸບ ແລະ ຂໍ້ສະເໜີ



## ສະຫຼຸບການຄົ້ນຄວ້າ

ໃນບົດນີ້ເປັນການສະຫຼຸບການຄົ້ນຄວ້າໃນຫົວຂໍ້ ການຮຽນຮູ້ ແລະ ການຈື່ຈຳ ຂໍ້ຄວາມຂຽນພາສາລາວໄປເປັນຂໍ້ຄວາມສຽງ ດ້ວຍເຕັກນິກຂອງປັນຍາປະດິດຂັ້ນສູງ. Text-To-Speech (TTS) ເປັນຂະບວນການທາງເອເລັກໂຕຣນິກ ໃນການແປຂໍ້ຄວາມຂຽນໄປເປັນຂໍ້ຄວາມສຽງໂດຍເຄື່ອງຄອມພິວເຕີ, ວິທີການນຳເອົາຂໍ້ຄວາມຂຽນ (Text) ເຂົ້າສູ່ລະບົບຄອມພິວຕີ ແມ່ນໄດ້ນໍາໃຊ້ຄີບອດ (Keyboard) ໃນການພິມຂໍ້ຄວາມໃສ່ໃນຊ່ອງພິມຂໍ້ຄວາມໃນເວັບໄຊ ເມື່ອພິມຂໍ້ຄວາມສຳເລັດແລ້ວ ເວັບໄຊຈະເຮັດການແປຂໍ້ຄວາມຂຽນໄປເປັນຂໍ້ຄວາມສຽງທັນທີທີ່ກົດປຸ່ມການແປຂໍ້ຄວາມຂຽນໄປເປັນຂໍ້ຄວາມສຽງ. ການແປຂໍ້ຄວາມຂຽນພາສາລາວຈາກຊ່ອງພິມຂໍ້ຄວາມໄປເປັນຂໍ້ຄວາມສຽງ ເຮັດໄດ້ຢ່າງວ່ອງໄວ ແລະ ຂ້ອນຂ້າງມີປະສິດທິພາບທີ່ດີ.

ການຮຽນຮູ້ຈື່ຈຳພາສາລາວຈາກສຽງ ແລະ ຕົວອັກສອນ ໂດຍນຳໃຊ້ສະຖາປັດຕະຍະ ກຳຂອງ DNN ແລະ RNN ເຊິ່ງແມ່ນ Algorithm ທີ່ຈັດຢູ່ໃນ Deep Learning ໂດຍຂັ້ນຕອນການດຳເນີນການນຳໃຊ້ ແບບຈຳລອງນັ້ນປະກອບໄປດ້ວຍ ການກະກຽມຂໍ້ມູນສຽງ (Dataset), ການກະກຽມກ່ອນການປະມວນຜົນ ແລະ ຂັ້ນຕອນການປະມວນຜົນ. ສຳລັບການກະກຽມຂໍ້ມູນແມ່ນໄດ້ເກັບກຳຈຳນວນຕົວອັກສອນພາສາລາວລວມທັງໝົດ 53 ຕົວ, ເຊິ່ງມີຟາຍສຽງພາສາລາວລວມທັງໝົດ 900 ຟາຍສຽງ. ສຳລັບການກະກຽມກ່ອນການປະມວນຜົນ ໄດ້ນຳຟາຍສຽງທັງໝົດໄປປ່ຽນເປັນຟາຍສຽງປະເພດ wav, ຕັດເອົາສະເພາະສຽງທີ່ເຮົາບັນທຶກ (ຫຼື ຕັດຊ່ອງວ່າງຂອງສຽງທີ່ບັນທຶກ), ຕັດສຽງລົບກວນອ້ອມຂ້າງ (Eliminate Noise) ແລະ ສໍາລັບຂັ້ນຕອນການປະມວນຜົນ ແມ່ນໄດ້ນຳໃຊ້ວິທີການຂອງ DNN ເພື່ອຮຽນຮູ້ສຽງພາສາລາວຈາກຟາຍສຽງ ເຊິ່ງໄດ້ນຳໃຊ້ໂຄງສ້າງ ສະຖາປັດຕະຍະກຳຂອງ RNN ນໍາມາປະມວນຜົນກໍຄື ແບບຈໍາລອງ YourTTS ທີ່ໄດ້ປັບປຸ່ງໂຄງສ້າງແລ້ວ. ໂດຍໄດ້ນໍາເອົາຂໍ້ມູນເພື່ອເຝິກສອນ (Train) ຈໍານວນ 855 ຟາຍສຽງ (ຂໍ້ມູນເຝິກສອນ 80%) ແລະ ຈໍານວນທີ່ນໍາມາທົດສອບ (Test) ມີ 45 ຟາຍສຽງ (ຂໍ້ມູນທົດສອບ 5%), ເຊິ່ງຜົນຂອງການຮຽນຮູ້ຂອງແບບຈໍາລອງ YourTTS ເຫັນວ່າມີ ຄ່າການເຝິກສອນໄດ້ຄິດເປັນຄ່າຕ່າງໆຄື: ເວລາສະເລ່ຍໃນການໂຫຼດຂໍ້ມູນ (avg\_loader\_time) ມີຄ່າ 0.0054 ວິນາທີ, ການສູນເສຍ mel-spectrogram (avg\_loss\_mel) ມີຄ່າ 20.29, ເວລາການສູນເສຍການຄາດຄະເນ ແລະ ໄລຍະເວລາຄວາມຈິງພື້ນຖານຂອງ phoneme ໃນການເວົ້າທີ່ສ້າງຂຶ້ນ (avg\_loss\_duration) ມີຄ່າ 0.34 ວິນາທີ ແລະ ການສູນເສຍ mel-spectrogram (avg\_loss\_1) ມີຄ່າ 30.27 ໃນການເຝິກສອນໃນຮອບທີ່ 100,000.

ການປະເມີນແບບຈຳລອງໂດຍການພະຍາກອນ ສຽງພາສາລາວຂອງແບບຈຳລອງ YourTTS ເຫັນວ່າສາມາດພະຍາກອນໄດ້ທຸກສຽງ ຍົກເວັ້ນແຕ່ ສຽງສັນຍາລັກອື່ນໆຄື: ໆ ທີ່ແບບຈຳລອງບໍ່ຮູ້ຈັກ. ສະນັ້ນ, ຈຶ່ງໄດ້ນຳເອົາແບບຈຳລອງນີ້ມາພັດທະນາເປັນເວັບໄຊ ໂດຍນຳໃຊ້ Flask Framework ເຂົ້າຊ່ວຍໃນການສ້າງເວັບໄຊ.

## ຂໍ້ຈຳກັດຂອງການຄົ້ນຄວ້າ

ຂໍ້ຈຳກັດໃນການສຶກສາຄົ້ນຄວ້າໃນຄັ້ງນີ້ມີຄື:

###### ການຮຽນຮູ້ຂອງແບບຈຳລອງ ຍັງບໍ່ສາມາດອ່ານຕົວເລກສາກົນໄດ້.

###### ການຮຽນຮູ້ຂອງແບບຈຳລອງ ຍັງບໍ່ສາມາດອ່ານຂໍ້ຄວາມທີ່ພິມໃນຊ່ອງພິມຂໍ້ຄວາມເວັບໄຊ ທີ່ບໍ່ມີສຽງໃນຊຸດຂໍ້ມູນ (Dataset) ທີ່ເຝິກສອນ (Train) ໄດ້.

###### ຂໍ້ຈຳກັດທາງດ້ານການອ່ານສັນຍາລັກ ຍັງບໍ່ສາມາດອ່ານສັນຍາລັກຕ່າງໆໄດ້.

###### ຄ່າການເຝິກສອນທີ່ຄິດເປັນຄ່າຕ່າງໆຍັງບໍ່ສາມາດຕໍ່າກວ່ານີ້ເຊັ່ນ: ເວລາສະເລ່ຍໃນການໂຫຼດຂໍ້ມູນ (avg\_loader\_time) ມີຄ່າ 0.0054 ວິນາທີ, ການສູນເສຍ mel-spectrogram (avg\_loss\_mel) ມີຄ່າ 20.29, ເວລາການສູນເສຍການຄາດຄະເນ ແລະ ໄລຍະເວລາຄວາມຈິງພື້ນຖານຂອງ phoneme ໃນການເວົ້າທີ່ສ້າງຂຶ້ນ (avg\_loss\_duration) ມີຄ່າ 0.34 ວິນາທີ ແລະ ການສູນເສຍ mel-spectrogram (avg\_loss\_1) ມີຄ່າ 30.27 ໃນການເຝິກສອນໃນຮອບທີ່ 100,000. ຄ່າທັງໝົດທີ່ກ່າວມາຍັງບໍ່ດີພໍ ເນື່ອງຈາກອຸປະກອນ ແລະ ສະພາບແວດລ້ອມທີ່ໃຊ້ໃນການບັນທຶກສຽງຍັງບໍ່ດີພໍ (ງົບປະມານບໍ່ພຽງພໍ).

## ຂໍ້ສະເໜີໃນການຄົ້ນຄວ້າຕໍ່ໄປ

ສຳລັບການຄົ້ນຄວ້າໃນຄັ້ງນີ້ແມ່ນໄດ້ນຳໃຊ້ເຕັກນິກຂອງ DNN ເຊິ່ງທີ່ເປັນໂຄງສ້າງ ສະຖາປັດຕະຍະກຳຂອງ RNN ເຂົ້າຊ່ວຍໃນການຈື່ຈໍາສຽງ ແລະ ນໍາໃຊ້ແບບຈໍາລອງ YourTTS ໃນການຄົ້ນຄວ້າກ່ຽວກັບການຮຽນຮູ້ສຽງ.

ການແນະນຳແນວທາງການພັດທະນາໄດ້ແກ່ ການພັດທະນາ Noise Removal ໃນສ່ວນຂອງ Pre-processing ເພື່ອໃຫ້ສາມາດລົບຄື້ນສຽງລົບກວນ ໃນກໍລະນີສຽງທີ່ມິຄື້ນສຽງລົບກວນຈຳນວນຫຼາຍອອກໄປ ແລະ ເພື່ອໃຫ້ແບບຈຳລອງສາມາດອ່ານຂໍ້ຄວາມສຽງໄດ້ຢ່າງມີປະສິດທິພາບຫຼາຍຂຶ້ນ.

ແນະນຳການຊອກຫາເຕັກນິກ ຫຼື Algorithm ກ່ຽວກັບ TTS ພາສາລາວ ທີ່ມີຄວາມຊັດເຈນ ແລະ ສາມາດປັບຄ່າຄວາມໄວ ແລະ ຊ້າຂອງສຽງເວລາປະມວນອອກມາ.

ແນະນຳການພັດທະນາໃນສ່ວນ Post-processing ທີ່ເປັນສ່ວນທີ່ສຳຄັນໃນການປັບພາສາລາວໃຫ້ຖືກຕ້ອງຕາມຫຼັກວັດຈະນານຸກົມພາສາລາວ.

ແນະນຳການພັດທະນາໃນສ່ວນ Training Data ໃຫ້ສາມາດສ້າງແບບຈໍາລອງກັບຊັບພະຍາກອນເຄື່ອງຂະໜາດນ້ອຍເຊັ່ນ: ເຄື່ອງຄອມພິວເຕີທີ່ມີ GPU ຕໍ່າ (Laptop ທີ່ປະປັດພະລັງງານເປັນຕົ້ນ) ຫຼື ບໍ່ມີ GPU ເລີຍ.

ສຳລັບການພັດທະນາໃນສ່ວນ STT ແມ່ນຢາກໃຫ້ພັດທະນາເພີ່ມເຕີມໃນສ່ວນທີ່ສາມາດແປງຂໍ້ຄວາມໄປເປັນຟາຍຕ່າງໆເຊັ່ນ: .txt, .docx, .doc, .pdf ແລະ ອື່ນໆ.

ເອກະສານອ້າງອີງ

**ເອກະສານອ້າງອີງ**

ກະຊວງໄປສະນີ ໂທລະຄົມມະນາຄົມ ແລະ ການສື່ສານ (2012). **ປະຫວັດພາສາລາວ**.

[On-line] Available: https://www.phetsarath.gov.la/gweb/backend/web/index.php?r=site/detail&id=472

ຜາຕຸງ (14/06/2023). ມີທາງດີທາງງາມ ຕ້ອງຮູ້ຮັກສາບົວລະບັດ ໃຫ້ສົມກັບຄຳວ່າ: ລັດກັບປະຊາຊົນຊ່ວຍ ກັນເຮັດ ຊ່ວຍກັນສ້າງສາພັດທະນາໃຫ້ຈະເລີນ. **ລາວພັດທະນາ**   
[On-line serial] Available: https://www.laophattananews.com/archives/164459

ຜາຕຸງ (12/06/2023). ຄວນກຳຈັດຮື້ຖອນອອກແຕ່ຫົວທີ ກ່ອນຈະມີນ້ຳຖ້ວມອັ່ງຂັງໃນຕົວເມືອງ. **ລາວພັດທະນາ** [On-line serial] Available: https://www.laophattananews.com/archives/164390

ຜາຕຸງ (10/06/2023). ກ້າວເຂົ້າສູ່ລະດູຝົນແລ້ວ ບໍ່ຢາກໃຫ້ນ້ຳຖ້ວມຂັງ ຕ້ອງຈິງຈັງຮ່ວມມືກັນແກ້ໄຂ. **ລາວພັດທະນາ** [On-line serial] Available: <https://www.laophattananews.com/archives/164232>

ຜາຕຸງ (09/06/2023). ປູກພືດກິນໄດ້​ແທນຢາສູບ” ເວົ້າຄືຊິງ່າຍ ແຕ່ເຮັດແທ້ໆຍັງມີຫລາຍປັດໄຈທີ່ຕ້ອງແກ້. **ລາວພັດທະນາ** [On-line serial] Available: https://www.laophattananews.com/archives/164132

ທິດກ່ຳ (08/06/2023). ຫລາຍຄົນລະດົມກັນປູກຕົ້ນໄມ້ໃຫ້ເປັນປ່າ ແຕ່ຢ່າໃຫ້ໃຜຄົນໃດໜຶ່ງໃຊ້ອາຍາສິດໃນການທຳລາຍ. **ລາວພັດທະນາ** [On-line serial] Available: https://www.laophattananews.com/archives/164003

ພູຊ້າງນ້ອຍ (07/06/2023). ການຖິ້ມຂີ້ເຫຍື້ອຊະຊາຍ ເປັນບັນຫາທີ່ຍັງແພ່ຫລາຍໃນສັງຄົມລາວ. **ລາວພັດທະນາ** [On-line serial] Available: https://www.laophattananews.com/archives/163905

ຜາຕຸງ (06/06/2023). ເອົາໃຈໃສ່ກະກຽມໃຫ້ດີ ແລະ ສອບເສັງໃຫ້ໄດ້ຕາມຄາດໝາຍຈຶ່ງບໍ່ຕົກໃສ່ຄຳວ່າ ເຮັດນາໝົດປີໄຟໄໝ້ເລົ້າເຂົ້າ. **ລາວພັດທະນາ** [On-line serial] Available: https://www.laophattananews.com/archives/163794

ຜາຕຸງ (02/06/2023). 4 ທ່າແຮງໃນການພັດທະເສດຖະກິດ ນໍາເອົາປະຊາຊົນເມືອງຊໍາເໜືອ ແຂວງຫົວພັນໃຫ້ຫລຸດພົ້ນອອກຈາກຄວາມທຸກຍາກ. **ລາວພັດທະນາ** [On-line serial] Available: https://www.laophattananews.com/archives/165830

ມຶກບໍ່ແຫ້ງ (02/06/2023). ການຂຸດຄົ້ນແຮ່ທາດ ທີ່ບໍ່ໄປຕາມມາດຕະຖານ ແລະ ກົດໝາຍ ກາຍເປັນບັນຫາເຄັ່ງຮ້ອນຂອງຊາດ. **ລາວພັດທະນາ** [On-line serial] Available: https://www.laophattananews.com/archives/165887

ຜາຕຸງ (06/06/2023). ຄວນຄຳນຶງເຖິງຄວາມປອດໄພກ່ອນຈະກິນອາຫານທີ່ປຸງແຕ່ງດ້ວຍເຫັດປ່າ. **ລາວພັດທະນາ** [On-line serial] Available: <https://www.laophattananews.com/archives/163812>

ຜາຕຸງ (12/05/2023). ກູ້ເງິນອອນລາຍງ່າຍ ບໍ່ມີຈິງ ລະວັງມິດສາຊີບລໍ້ລວງເອົາຂໍ້ມູນສ່ວນຕົວ. **ລາວພັດທະນາ** [On-line serial] Available: <https://www.laophattananews.com/archives/161321>

ທິດກ່ຳ (11/05/2023). ບານເຕະທີມຊາດລາວໄປບໍ່ເຖິງ ຕາມທີ່ປວງຊົນລາວຕັ້ງເປົ້າໝາຍ ເຊິ່ງມີຫລາຍຢ່າງຕ້ອງພັດທະນາ. **ລາວພັດທະນາ** [On-line serial] Available: <https://www.laophattananews.com/archives/161241>

ທິດກ່ຳ (10/05/2023). ທາງການສັ່ງປິດຮ້ານແລກປ່ຽນເງິນຕາແລ້ວກໍຈິງ ແຕ່ຍັງປາກົດມີການລັກແລກປ່ຽນນອກລະບົບແບບຜິດກົດໝາຍຢູ່. **ລາວພັດທະນາ** [On-line serial] Available: <https://www.laophattananews.com/archives/161036>

ທິດກ່ຳ (08/05/2023). ຍຸກຂໍ້ມູນຂ່າວສານທັນສະໄໝ ຢ່າໃຫ້ຕົກເປັນເຫຍື່ອກຸ່ມຕົ້ມຕຸນຂາຍຝັນລໍ້ລວງໃຫ້ເກີດຄວາມໂລບ. **ລາວພັດທະນາ** [On-line serial] Available: <https://www.laophattananews.com/archives/160940>

ພານິດາ ພານິດຊະກຸນ (2009). **ເຕັກໂນໂລຢີເຊີງວັດຖຸ (Object-Oriented technology)**

IT-HR (2014). **ການຂຽນແຜນງານໂດຍ Microsoft Visio 2007**. [On-line] Available: https://www.rama.mahidol.ac.th/hr/sites/default/files/public/img/Gallerry/HR-02/pdf\_file/All.pdf

Sangramsing N. Kayte. (2015) **Marathi Text-To-Speech Synthesis using Natural Language Processing**. [On-line] Available: <https://www.researchgate.net/publication/284294625_Marathi_Text-To-Speech_Synthesis_using_Natural_Language_Processing>

Chowdhury, G. G. (2003). **Natural language processing. Annual Review of Information Science and Technology**, 37(1), 51-89.

Aone, C., Okurowski, M. E., & Gorlinsky, J. (1998, August). Trainable, scalable summarization using robust NLP and machine learning. **In Proceedings of the 17th International Conference on Computational Linguistics-Volume 1** (pp. 62-66). Association for Computational Linguistics.

Schank, R. C., & Abelson, R. P. (2013). **Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structures**. Psychology Press.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). **Deep Learning. MIT** press.

Hecht-Nielsen, R. (1992). Theory of the backpropagation neural network. **In Neural Networks for Perception** (pp. 65-93). Academic Press.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. **Nature**, 521(7553), 436-444.

Han, S., Liu, X., Mao, H., Pu, J., Pedram, A., Horowitz, M. A., & Dally, W. J. (2016). EIE: efficient inference engine on compressed deep neural network. **ACM SIGARCH Computer Architecture News**, 44(3), 243-254

Mikolov, T., Karafiát, M., Burget, L., Černocký, J., & Khudanpur, S. (2010). Recurrent neural network based language model. **In Proceeding of Eleventh Annual Conference of the International Speech Communication Association**.

Premjith, B., Soman, K. P., & Kumar, M. A. (2018). A deep learning approach for Malayalam morphological analysis at character level. **Procedia Computer Science**, 132, 47-54.

Casanova, Weber, Shulby, Junior, Golge and Antonelli, (2023). **YourTTS: Towards Zero-Shot Multi-Speaker TTS and Zero-Shot Voice Conversion for everyone**.[On-line] Available: <https://arxiv.org/pdf/2112.02418.pdf>

Kim, Kong and Son, (2021). **Conditional Variational Autoencoder with Adversarial Learning for End-to-End Text-to-Speech**. [On-line] Available: <https://arxiv.org/pdf/2106.06103.pdf>

Google Cloud, (2023-05-26)**. Speech-to-Text request construction**. [On-line] Available: [https://cloud.google.com/speech-to-text/docs/speech-to-text-requests#streaming-recognition](https://cloud.google.com/speech-to-text/docs/speech-to-text-requests%23streaming-recognition)

ເອກະສານຊ້ອນທ້າຍ

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ລ/ດ | ໜ້າວຽກທີ່ດຳເນີນງານ | ເວລາທີ່ໃຊ້ | ໄລຍະເວລາ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2022 | | | | 2023 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| ທັນວາ | | | | ມັງກອນ | | | | ກຸມພາ | | | | ມີນາ | | | | ເມສາ | | | | ພຶດສະພາ | | | | ມີຖຸນາ | | | | ກໍລະກົດ | | | | ສິງຫາ | | | |
| 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | |
| 1 | ຈັດຕັ້ງທີມ | 1 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 2 | ກຳນົດຫົວຂໍ້ | 1 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 3 | ລົງມືຂຽນບົດສະເໜີໂຄງການ | 2 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 4 | ກຳນົດບັນຫາ ແລະ ຂອບເຂດ | 1 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 5 | ສົ່ງບົດສະເໜີໂຄງການ | 1 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 6 | ປ້ອງກັນບົດສະເໜີໂຄງການ | 1 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 7 | ລົງມືຂຽນບົດຈົບຊັ້ນ | 22 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 8 | ສຶກສາທິດສະດີທີ່ກ່ຽວຂ້ອງ | 20 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 9 | ຮວບຮວມ ແລະ ສຶກສາຂໍ້ມູນ | 12 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 10 | ກະກຽມຂໍ້ມູນ | 20 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 11 | ສ້າງແບບຈຳລອງຂໍ້ມູນ | 8 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 12 | ປະເມີນຜົນແບບຈຳລອງ | 2 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 13 | ນຳໄປໃຊ້ງານ | 4 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 14 | ອອກແບບໜ້າຕາເວັບໄຊ | 1 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 15 | ລົງມືຂຽນເວັບໄຊ | 2 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 16 | ທົດສອບເວັບໄຊ | 2 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 17 | ແກ້ໄຂຂໍ້ຜິດພາດຂອງເວັບໄຊ | 4 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 18 | ສົ່ງບົດໃຫ້ຜູ້ຊ່ວຍນຳພາ | 1 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 19 | ກວດແກ້ບົດ | 2 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 20 | ສ້າງຄູ່ມືນຳໃຊ້ເວັບໄຊ | 1 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 21 | ຊ້ອມປ້ອງກັນບົດ | 3 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 22 | ປ້ອງກັນບົດ | 1 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 23 | ແກ້ໄຂບົດ | 2 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 24 | ສົ່ງບົດສົມບູນ | 1 ອາທິດ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
| 25 | ລວມເວລາທີ່ໃຊ້ | 36 ອາທິດ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

**ໄລຍະເວລາການສຶກສາ**

ປະຫວັດຫຍໍ້ຜູ້ຂຽນບົດ

**ປະຫວັດຫຍໍ້ຜູ້ຂຽນບົດ**

ຊື່ ແລະ ນາມສະກຸນ: ທ. ຄຳປະເສີດ ໄຊຍະວົງ

ວັນທີ, ເດືອນ, ປີເກີດ: 22 ກຸມພາ 2002

ບ້ານເກີດ: ບ້ານ ນາສ້ຽວ, ເມືອງ ນາຊາຍທອງ, ນະຄອນຫຼວງວຽງຈັນ

ປັດຈຸບັນ: ບ້ານ ນາສ້ຽວ, ເມືອງ ນາຊາຍທອງ, ນະຄອນຫຼວງວຽງຈັນ

ການສຶກສາ:

ປີ 2019-2023 ໄດ້ເປັນນັກສຶກສາໃນລະດັບປະລິນຍາຕີຢູ່ທີ່ ມະຫາວິທະຍາໄລແຫ່ງຊາດ

ປີ 2015-2019 ໄດ້ຈົບການຮຽນຊັ້ນມັດທະຍົມສຶກສາຕອນປາຍຢູ່ທີ່ ໂຮງຮຽນມັດທະຍົມສຶກສາສົມບູນຈັນສະຫວ່າງ

ປີ 2012-2015 ໄດ້ຈົບການຮຽນຊັ້ນມັດທະຍົມສຶກສາຕອນຕົ້ນຢູ່ທີ່ ໂຮງຮຽນມັດທະຍົມສຶກສາຕອນຕົ້ນໜອງທອງ

ປີ 2007-2012 ໄດ້ຈົບການຮຽນຊັ້ນປະຖົມສົມບູນຢູ່ທີ່ ໂຮງຮຽນປະຖົມສົມບູນລີນດ້າ

ເບີໂທ ແລະ WhatsApp: (+85620) 5457-3377, (+85620) 9587-4963

Email: 205n001419@nuol.edu.la

**ປະຫວັດຫຍໍ້ຜູ້ຂຽນບົດ**

ຊື່ ແລະ ນາມສະກຸນ: ທ້າວ ໃຫຍ່ ແສງວິໄຊ

ວັນທີ, ເດືອນ, ປີເກີດ: 17 ເມສາ 2000

ບ້ານເກີດ: ບ້ານ ຫາດສວນ ເມືອງ ທຸລະຄົມ ແຂວງ ວຽງຈັນ

ປັດຈຸບັນ: ບ້ານ ຫາດສວນ ເມືອງ ທຸລະຄົມ ແຂວງ ວຽງຈັນ

ການສຶກສາ:

ປີ 2019-2023 ໄດ້ເປັນນັກສຶກສາໃນລະດັບປະລິນຍາຕີຢູ່ທີ່ ມະຫາວິທະຍາໄລແຫ່ງຊາດ

ປີ 2011-2018 ໄດ້ຈົບການຮຽນຊັ້ນມັດທະຍົມຢູ່ທີ່ ໂຮງຮຽນມັດທະຍົມສຶກສາສົມບູນເວີນຄໍາ

ປີ 2006-2011 ໄດ້ຈົບການຮຽນຊັ້ນປະຖົມສົມບູນຢູ່ທີ່ ໂຮງຮຽນປະຖົມສົມບູນຫາດສວນ

ເບີໂທ ແລະ WhatsApp: (+85620) 59933659, (+85620) 58578322

Email: gnaysaengvixai1740@gmail.com