



ສາທາລະນະລັດ ປະຊາທິປະໄຕ ປະຊາຊົນລາວ
ສັນຕິພາບ ເອກະລາດ ປະຊາທິປະໄຕ ເອກະພາບ ວັດທະນາຖາວອນ
-----((0))-----



ມະຫາວິທະຍາໄລແຫ່ງຊາດ
ຄະນະວິທະຍາສາດທຳມະຊາດ
ພາກວິຊາວິທະຍາສາດຄອມພິວເຕີ

ບົດສະເໜີໂຄງການຈົບຊັ້ນ ລະດັບປະລິນຍາຕີວິທະຍາສາດ
ສາຂາ ວິທະຍາສາດຄອມພິວເຕີ

ຊື່ບົດຈົບຊັ້ນ (Title)

ພາສາລາວ: ເວັບໄຊວັດຈະນານຸກົມສຽງເວົ້າພາສາລາວ-ອັງກິດ
ພາສາອັງກິດ: Lao - English Speech Dictionary Website

ລາຍຊື່ນັກສຶກສາຂຽນບົດ (Project Team)

ລ/ດ	ລະຫັດ	ຊື່ ແລະ ນາມສະກຸນ	ເບີໂທ
1	205N0014.19	ທ. ຄຳປະເສີດ ໄຊຍະວົງ	02054573377
2	205N0045.19	ທ. ໃຫຍ່ ແສງວິໄຊ	02059933659

ອາຈານທີ່ປຶກສາ (Advisor)

ລ/ດ	ຊື່ ແລະ ນາມສະກຸນ	ຕຳແໜ່ງ
1		ນຳພາ
2		ຊ່ວຍນຳພາ

1. ຄວາມສໍາຄັນຂອງບັນຫາ

ໃນປັດຈຸບັນໂລກຂອງພວກເຮົາໄດ້ມີການພັດທະນາຄວາມສາມາດຂອງອຸປະກອນຕ່າງໆ ດ້ານຄອມພິວເຕີ ໃຫ້ມີການຮຽນຮູ້ຕາມແບບຢ່າງຂອງມະນຸດ. ໃນດ້ານການຈື່ຈໍາສຽງເວົ້າ (Speech Recognition) ກໍມີຄວາມກ້າວໜ້າໄປຫຼາຍ ໄດ້ມີການນໍາໄປປະຍຸກໃຊ້ໃນລະບົບສື່ສານໂທລະຄົມມະນາຄົມ ເຊັ່ນ: ການສັ່ງໂທອອກຂອງໂທລະສັບມືຖືໂດຍໃຊ້ສຽງ. ຈຸດປະສົງຫຼັກຂອງການຈື່ຈໍາສຽງເວົ້າຄືການເພີ່ມຄວາມສາມາດໃຫ້ອຸປະກອນຕ່າງໆ ສາມາດຮັບຮູ້ ແລະ ມີການຕອບກັບມະນຸດໄດ້ຫຼາຍຂຶ້ນ ເຊິ່ງການໃຊ້ສຽງເວົ້າໃນການສື່ສານຄວບຄຸມສັງການຖືວ່າເປັນວິທີທີ່ມະນຸດໃຊ້ຢ່າງເປັນທໍາມະຊາດທີ່ສຸດ. ເນື່ອງຈາກ ພາສາລາວຍັງບໍ່ມີການສ້າງລະບົບການແປຂໍ້ຄວາມພາສາລາວມາເປັນສຽງເວົ້າ (Text To Speech) ທີ່ສົມບູນ ແລະ ຖືກຕ້ອງທີ່ສຸດ. ພວກຂ້າພະເຈົ້າ ຈຶ່ງຢາກສຶກສາລະບົບການແປຂໍ້ຄວາມພາສາລາວມາເປັນສຽງເວົ້າ (Text To Speech) ໃຫ້ສົມບູນ ແລະ ຖືກຕ້ອງທີ່ສຸດ.

ວັດຈະນານຸກົມ (ຄໍາເຄົ້າ: ວັຈນານຸກົມ, ອັງກິດ: dictionary, ຝຣັ່ງ: dictionnaire) ແມ່ນປຶ້ມອ້າງອີງປະເພດໜຶ່ງໂດຍທົ່ວໄປ ໝາຍເຖິງປຶ້ມທີ່ຮິບໂຮມຄໍາສັບໃນວົງສັບທີ່ກຳໜົດ ແລະ ນິຍາມຄວາມໝາຍເອົາໄວ້ ເພື່ອໃຊ້ເປັນທີ່ຊອກຫາຄວາມໝາຍຂອງຄໍາ ໂດຍມີການຮຽງລຳດັບຄໍາສັບຕາມໂຕອັກສອນ, ຕາມສຽງ ຫຼື ຕາມລຳດັບອື່ນໆ ທີ່ເໝາະສົມສອດຄ້ອງກັບການໃຊ້ວັດຈະນານຸກົມນັ້ນໆ. ວັດຈະນານຸກົມຍັງໝາຍເຖິງ ປຶ້ມທີ່ໃຫ້ລາຍລະອຽດ ຄອບຄຸມວົງສັບທີ່ກວ້າງ ຂະນະທີ່ປຶ້ມຮິບໂຮມ ແລະ ອະທິບາຍຄໍາສັບໃນວົງແຄບ ແລະ ມີຈຳນວນຈຳກັດ ມັກຈະເອີ້ນວ່າ: ປະທານຸກົມ ຢ່າງໃດກໍດີ ຄຳວ່າປະທານຸກົມ ແລະ ວັດຈະນານຸກົມອາດໃຊ້ສະຫຼັບກັນໄດ້. ເນື່ອງຈາກປັດຈຸບັນໂລກຂອງເຮົາກໍໄດ້ກ້າວສູ່ຍຸກຂອງດິຈິຕອລ (Digital) ທີ່ທຸກໆຄົນກໍຕ່າງມີອຸປະກອນຄອມພິວເຕີ (Computer) ແລະ ສະມາດໂຟນ (Smart Phone) ເປັນຂອງຕົນເອງ ໃນການນໍາໃຊ້ຊີວິດປະຈຳວັນ. ໂດຍສະເພາະການອ່ານນະວະນິຍາຍ (Novel), ປຶ້ມແບບເອັກເລັກໂຕຣນິກ (E-Book), ການຮັບຊົມຮູບເງົາ, ຝັງເພງຕ່າງກໍຜ່ານທາງອຸປະກອນເຫຼົ່ານີ້, ເຊິ່ງວັດຈະນານຸກົມເອງກໍເຊັ່ນດຽວກັນ ບໍ່ວ່າຈະເປັນ ປຶ້ມວັດຈະນານຸກົມຂອງມະຫາວິທະຍາໄລຊື່ດັງຂອງຕ່າງປະເທດເຊັ່ນວ່າ: Oxford Dictionary, Cambridge Dictionary, ແອັບພິເຄເຊັນແປພາສາອອນລາຍບໍ່ວ່າຈະເປັນ Google Translate, Microsoft Translator ແລະ ອື່ນໆ.

ສະນັ້ນ, ພວກຂ້າພະເຈົ້າ ຈຶ່ງມີຄວາມສົນໃຈຢາກຈະພັດທະນາ ເວັບໄຊວັດຈະນານຸກົມສຽງເວົ້າພາສາລາວ-ອັງກິດ (Lao - English Speech Dictionary Website) ເພື່ອເຂົ້າມາຊ່ວຍໃນການສື່ສານ-ສົນທະນາລະຫວ່າງພາສາລາວ-ອັງກິດໃຫ້ມີຄວາມສະດວກສະບາຍ ແລະ ການສຶກສາ-ຮຽນຮູ້ພາສາລາວ-ອັງກິດ ໃຫ້ມີຄວາມຖືກຕ້ອງທີ່ສຸດ.

2. ຈຸດປະສົງຂອງການຄົ້ນຄ້ວາ (Objectives)

- ເພື່ອສຶກສາຂະບວນການເຮັດວຽກຂອງ ລະບົບການຈື່ຈໍາສຽງເວົ້າ (Speech Recognition)
- ເພື່ອສ້າງເວັບໄຊວັດຈະນານຸກົມສຽງເວົ້າພາສາລາວ-ອັງກິດ (Lao - English Speech Dictionary Website) ແບບ Client-Server
- ເພື່ອໃຫ້ມີຄວາມສະດວກສະບາຍ ແລະ ຖືກຕ້ອງທີ່ສຸດໃນການສື່ສານ-ສົນທະນາລະຫວ່າງພາສາລາວ-ອັງກິດ
- ເພື່ອສຶກສາ-ຮຽນຮູ້ ແລະ ເພີ່ມທັກສະທາງດ້ານພາສາລາວ-ອັງກິດແກ່ຜູ້ໃຊ້ງານ

3. ຂອບເຂດການຄົ້ນຄວ້າ (Scope)

ເວັບໄຊວັດຈະນານຸກົມສຽງເວົ້າພາສາລາວ-ອັງກິດ (Lao - English Speech Dictionary Website) ນີ້ໄດ້ສຶກສາມາຈາກບັນຫາໃນການສື່ສານ-ສົນທະນາແບບບ້ອນຂໍ້ມູນໄປເປັນແບບການຈື່ຈຳສຽງເວົ້າ (Speech Recognition) ທີ່ມີຂອບເຂດໃນການເຮັດວຽກດັ່ງນີ້:

- ສຶກສາ ແລະ ນຳໃຊ້ແບບຈຳລອງຂໍ້ມູນທີ່ຖືກຜັດທະນາຂຶ້ນໂດຍ Google Speech To Text API ເພື່ອແປຄຳເວົ້າພາສາລາວມາເປັນຂໍ້ຄວາມ (Lao Speech To Text)
- ສ້າງຂັ້ນຕອນວິທີແບບ Deep Learning Neural Network ໂດຍນຳໃຊ້ວິທີການຮຽນຮູ້ແບບການຈື່ຈຳສຽງເວົ້າ (Speech Recognition) ເພື່ອສ້າງແບບຈຳລອງການແປຂໍ້ຄວາມມາເປັນສຽງເວົ້າພາສາລາວ (Lao Text To Speech)
- ສ້າງເວັບໄຊວັດຈະນານຸກົມສຽງເວົ້າພາສາລາວ-ອັງກິດຈາກແບບຈຳລອງຂໍ້ມູນທີ່ເໝາະສົມທີ່ສຸດ

4. ປະໂຫຍດຄາດວ່າຈະໄດ້ (Expected Outcome of the Project)

ໄດ້ຮູ້ເຖິງຂະບວນການເຮັດວຽກຂອງ ລະບົບການຈື່ຈຳສຽງເວົ້າ (Speech Recognition) ແລະ ບັນຫາທີ່ເກີດຂຶ້ນຂອງລະບົບຈື່ຈຳສຽງເວົ້າພາສາລາວ.

- ໄດ້ເວັບໄຊວັດຈະນານຸກົມສຽງເວົ້າພາສາລາວ-ອັງກິດ (Lao - English Speech Dictionary Website) ໃໝ່.
- ໄດ້ເວັບໄຊທີ່ມີຄວາມສະດວກສະບາຍ ແລະ ຖືກຕ້ອງທີ່ສຸດໃນການສື່ສານ-ສົນທະນາລະຫວ່າງພາສາລາວ-ອັງກິດ.
- ສາມາດສຶກສາ-ຮຽນຮູ້ ແລະ ເພີ່ມທັກສະທາງດ້ານພາສາອັງກິດ ຈາກເວັບໄຊວັດຈະນານຸກົມສຽງເວົ້າພາສາລາວ-ອັງກິດ (Lao - English Speech Dictionary Website) ຂອງພວກຂ້າພະເຈົ້າໄດ້.

5. ວິທີດຳເນີນການຄົ້ນຄວ້າ (Research Methodology)

ຂັ້ນຕອນ ແລະ ການດຳເນີນໂຄງການໂດຍຫຼັກໆແລ້ວ ພວກຂ້າພະເຈົ້າໄດ້ອີງໃສ່ຕາມຂະບວນການ Natural language processing (NLP) ດັ່ງລຸ່ມນີ້:

5.1 ຂັ້ນຕອນກ່ອນການປະມວນຜົນ (Preprocessing)

ໃນໄລຍະຂັ້ນຕອນກ່ອນການປະມວນຜົນເພື່ອກະກຽມຂໍ້ຄວາມດິບໃຫ້ເໝາະສົມເພື່ອເຂົ້າສູ່ຂັ້ນຕອນການປະມວນຜົນ, ກວດຈັບຄຳໃນແຕ່ລະປະໂຫຍກ, ຊ່ອງວ່າງໃນແຕ່ລະຄຳ, ເຄື່ອງໝາຍວັກຕອນ ຫຼື ສັນຍາລັກອື່ນໃດທີ່ບໍ່ແມ່ນພາສາລາວ.

5.2 ການປະມວນຜົນພາສາທຳມະຊາດ (Natural Language Processing)

ຂັ້ນຕອນນີ້ຂຶ້ນຢູ່ກັບຖານຂໍ້ມູນທີ່ຜັດທະນາເພື່ອປຽບທຽບແຕ່ລະຄຳກັບສຽງເວົ້າທີ່ທຽບເທົ່າກັນເຊິ່ງປະກອບດ້ວຍ 3 ສ່ວນຫຼັກໆດັ່ງນີ້:

- ຂໍ້ຍົກເວັ້ນຂອງວັດຈະນານຸກົມ: ປະກອບມີລາຍການຄຳທີ່ມີການອອກສຽງຢ່າງຊັດເຈນ ແທນທີ່ຈະຖືກກຳນົດດ້ວຍກົດການອອກສຽງ ຖ້າວ່າໃຊ້ກົດການອອກສຽງກັບຄຳ ແລະ ລະບຸຄຳນຳໜ້າ

ຫຼື ຄຳຕໍ່ທ້າຍມາດຕະຖານ ຄຳນັ້ນເອງຈະຖືກຄົ້ນຫາອີກຄັ້ງໜຶ່ງໃນລາຍການວັດຈະນານຸກົມ ຍົກເວັ້ນ ຫຼັງຈາກທີ່ຄຳນຳໜ້າ ຫຼື ຄຳຕໍ່ທ້າຍຖືກຕັດຖິ້ມໄປ

- ກົດການອອກສຽງຂອງພາສາ (ຕົວອັກສອນເປັນສຽງ): ສ່ວນນີ້ຂອງຖານຂໍ້ມູນຈະລະບຸໜ່ວຍສຽງທີ່ໃຊ້ໃນການອອກສຽງແຕ່ລະຕົວ ຫຼືວ່າ ລຳດັບຂອງຕົວອັກສອນ ກົດບາງຂໍ້ກໍຈະໃຊ້ສະເພາະເມື່ອຕົວອັກສອນນຳໜ້າ ຫຼື ຕາມດ້ວຍຕົວອັກສອນອື່ນໆ.
- ໜ່ວຍສຽງຂອງພາສາ: ກຳໜົດໜ່ວຍສຽງທັງໝົດທີ່ໃຊ້ໃນພາສາ ພ້ອມກັບສະແດງລັກສະນະແລະ ຂໍ້ມູນທີ່ຈະຜະລິດເປັນສຽງ, ສຽງສະຫຼະຖືກສ້າງຂຶ້ນຈາກການສັງເຄາະສຽງໃນຂະນະທີ່ພະຍັນຊະນະຟີເສດ ຈະຖືກເກັບບັນທຶກເປັນໄຟລ໌ Wav ລ່ວງໜ້າຕາມການສັງເຄາະສຽງແບບຮຽງຕໍ່ກັນ

5.3 ຂັ້ນຕອນການແປງອອກມາເປັນສັນຍານດິຈິຕອລ (Digital Signal Processing)

ໃນສ່ວນນີ້ຂອງລະບົບສັງເຄາະສຽງ ການແທນຄ່າໜ່ວຍສຽງທີ່ເປັນຜົນລັບຂອງຂໍ້ຄວາມທີ່ປ້ອນໂດຍການເນັ້ນສຽງຟີເສດ ແລະ ຢຸດແບບເປັນຈັງຫວະຈະຖືກປ່ຽນເປັນສຽງເວົ້າທີ່ເໝາະສົມ ແລະ ສາມາດບັນທຶກເປັນຝາຍຮູບຄືນສຽງໄດ້

6. ສະຖານທີ່ສຶກສາ (Study Site)

ສະຖານທີ່ການສຶກສາ ພວກຂ້າພະເຈົ້າ ໄດ້ດຳເນີນການຄົ້ນຄວ້າຢູ່ທີ່ ພາກວິຊາວິທະຍາສາດຄອມພິວເຕີ, ຄະນະວິທະຍາສາດທຳມະຊາດ, ມະຫາວິທະຍາໄລແຫ່ງຊາດ.

7. ໄລຍະປະຕິບັດ (Duration)

ແຜນການດຳເນີນງານຂອງ "ເວັບໄຊວັດຈະນານຸກົມສຽງເວົ້າພາສາລາວ-ອັງກິດ" ໂດຍເລີ່ມການສຶກສາຄວາມເປັນໄປໄດ້ໃນການສຶກສາດັ່ງຕາຕະລາງລຸ່ມນີ້:

			ໄລຍະເວລາ																															
			2022				2023																											
			ທັນວາ				ມັງກອນ				ກຸມພາ				ມີນາ				ເມສາ				ພຶດສະພາ				ມິຖຸນາ				ກໍລະກົດ			
ລ/ດ	ໜ້າວຽກທີ່ດຳເນີນງານ	ເວລາທີ່ໃຊ້	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	9 ອາທິດ																																
9	ຮວບຮວມ ແລະ ສຶກສາຂໍ້ມູນ	2 ອາທິດ																																
10	ກະກຽມຂໍ້ມູນ	8 ອາທິດ																																
11	ສ້າງແບບຈຳລອງຂໍ້ມູນ	4 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	1 ອາທິດ																																
20	ສ້າງຄູ່ມືນຳໃຊ້ເວັບໄຊ	1 ອາທິດ																																
21	ຊ້ອມປ້ອງກັນບົດ	1 ອາທິດ																																
22	ປ້ອງກັນບົດ	1 ອາທິດ																																
23	ແກ້ໄຂບົດ	2 ອາທິດ																																
24	ສົ່ງບົດສົມບູນ	1 ອາທິດ																																
25	ລວມເວລາທີ່ໃຊ້		32 ອາທິດ																															

32 ອາທິດ

8. ເຄື່ອງມືທີ່ນຳໃຊ້ໃນການພັດທະນາ (Development Tools)

❖ Hardware

- ເຄື່ອງທີ່ 1 ໃຊ້ໃນການສ້າງແບບຈຳລອງ ແລະ Server:
 - ASUS ROG Strix G531GU-AL001T
 - Processor Intel® Core™ I7-9th Generation
 - Nvidia GeForce GTX 1660Ti Graphic Card
 - 24GB RAM
 - Operation: Microsoft Windows, Ubuntu
- ເຄື່ອງທີ່ 2 ໃຊ້ໃນການອອກແບບໜ້າຕາ ແລະ ກະກຽມເອກະສານ:
 - Lenovo Ideapad G40-70
 - Processor Intel® Core™ I5-4th Generation
 - Intel® HD Graphic Family Graphic Card
 - 8GB RAM
 - Operation: Microsoft Windows

❖ Software

- IDE: Microsoft Visual Studio Code
- Framework and Library: Flutter SDK, Google Text to Speech API, TensorFlow
- Design and Drawing: Figma, Microsoft Visio 2019
- Programming Languages: Python, Dart, SQL
- Database: Oracle MySQL Server
- Documentation: Microsoft Office 365
- Server: Nginx

9. ເອກະສານອ້າງອີງ (Reference)

ສົມຈິຕ ພັນລັກ. (2012) ພາສາລາວລ້ານຊ້າງ ກ່ອນປີ ພ.ສ 2478; ຄ.ສ 1935 ສະບັບຄົ້ນຄວ້າ. ສົມມະນາ ການພິມ ສປປ ລາວ.

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

Pangu IOT. (2019) **ວິທີໃຊ້ Text- to -Speech ຂອງ Google**. [On-line] Available:
<https://medium.com/@pangu2158/%E0%B8%A7%E0%B8%B4%E0%B8%98%E0%B8%B5%E0%B8%81%E0%B8%B2%E0%B8%A3%E0%B9%83%E0%B8%8A%E0%B9%89-text-to-speech-%E0%B8%82%E0%B8%AD%E0%B8%87-google-81d0ac00a7e3>

ວັນທີ/...../.....

ລາຍເຊັນຄະນະກຳມະການ

ວັນທີ/...../.....

ລາຍເຊັນອາຈານທີ່ປຶກສາ

ວັນທີ/...../.....

ລາຍເຊັນນັກສຶກສາ