



ДУУ, ДҮРС БИЧЛЭГ, ЯРИАНЫ АНГЛИ-МОНГОЛ МАШИН ОРЧУУЛГА

Б.Бат-Эрдэнэ⁷

Abstract: In modern times, thanks to technological developments, the use of machine translation has simplified the work methods of translators, saved time, and increased labor productivity. The process of translating audio, video, and speech from the source language to the target language used to take a lot of time and manpower to translate using traditional translation methods, but the use of machine translation has created new advances in the field of translation studies, combining theory and practice to make realistic calculations and produce highly probable machine translation results. However, the quality of translation directly depends on the quality of the input sentence, and first the English voice needs to be recognized and converted to text with high quality. After that, it is translated into the target language, and finally converted into speech in the target language. The main purpose of this research work is to conduct basic research on the development of English-Mongolian machine translation software that translates audio, video, and speech in real time.

Түлхүүр үг: англи-монгол орчуулга, яриа таних, бичвэрийг яриа болгон хөрвүүлэх, нейрон сүлжээнд суурилсан машины орчуулга

1. Удиртгал

Дараах судлаачдын судалгааны бүтээл нь эх хэл боловсруулах хийгээд машин сургалтын салбарт гүн гүнзгий нөлөө үзүүлж, орчин үеийн хэлний загвар, алгоритмын хөгжлийг бий болгосон.

Bahdanau et al. (2015) нейрон сүлжээнд суурилсан машин орчуулгад анхаарлын механизмын тухай ойлголтыг танилцуулж, загваруудад оролт, гаралтын дарааллыг илүү үр дүнтэйгээр тохируулах боломжийг олгосон. Энэхүү нээлт нь машин орчуулгын системд хувьсгал хийж, илүү үнэн зөв, контекст хамааралтай орчуулгыг хөгжүүлэхэд хүргэсэн. Тэдний ажил нь NLP-ийн seq2seq загваруудын дараагийн ахиц дэвшлийн үндэс суурийг тавьсан юм.

Vaswani et al. (2017) Хувирган өөрчлөх архитектурыг нэвтрүүлсэн бөгөөд энэ нь үгсийн хоорондын харилцааг дарааллаар нь олж авахын тулд өөртөө анхаарал хандуулах механизмд тулгуурладаг. Энэхүү архитектур нь NLP-ийн янз бүрийн даалгавруудад хамгийн сүүлийн үеийн гүйцэтгэлд хүрч, олон тохиолдолд давтагдах нейрон сүлжээ (RNN) болон конволюцийн нейрон сүлжээ (CNN) -ний гүйцэтгэлээс давж гарсан.

Peters et al. (2018) үгийн контекст хамааралтай утгыг агуулсан үг оруулах аргыг бий болгосон. Тэдний судалгааны гол давуу тал их хэмжээний бичвэр өгөгдлүүд дээр бэлтгэгдсэн хоёр чиглэлтэй урт богино хугацааны санах ой ашигладагт оршино.

⁷ ХУИС, доктор

Wang et al. (2019) хувирган өөрчлөх архитектурыг машин орчуулгад зориулж гүйцэтгэх чадварыг сайжруулах арга техникийг санал болгосон. Тэдний ажил нь илүү үр дүнтэй сургалтын алгоритмууд болон орчуулгын нарийн төвөгтэй даалгавруудыг илүү нарийвчлалтай, хурдтайгаар шийдвэрлэх чадвартай ба нүсэр хэлний загваруудыг хөгжүүлэхэд хувь нэмэр оруулсан.

OpenAI компанийн бүтээсэн хиймэл оюун ухаантай харилцан ярианы программ болох ChatGPT (OpenAI, 2023), Anthropic компанийн Claude (Aman, 2023), Google компанийн хөгжүүлсэн Gemini (Gemini team and Google, 2024), Meta код болон сургагдсан загварыг нээлттэй эх сурвалжтайгаар байршуулсан Llama (Llama Team and Meta, 2023), мөн нээлттэй эх сурвалжтайгаар байршуулсан DeepSeek (DeepSeek AI, 2025) зэрэг LLM хүнтэй төстэй харилцан ярианы мэдрэмжийг хэрэглэгчдэд хүргэхийн тулд эх хэлний боловсруулалтыг маш өндөр түвшинд хийдэг болжээ.

Batsukh, BE (2023) өгүүлбэрийн бүтэц, хүрээг нейрон сүлжээнд суурилсан машины орчуулгын системд хэрхэн сургаж болох талаар өгүүлсэн.

Alec Radford болон түүний нөхөд 2022 оны судалгаандаа яриа танилтын хялбаршуулсан аргыг танилцуулсан бол Hideyuki Tachibana нөхдийн хамтаар 2017 онд яриа бичвэр хоёрыг ашиглан сургах боломжтой бичвэрийг ярианд хөрвүүлдэг аргачлалыг нээсэн.

Судлаачид цаашид эдгээр суурь судалгаанууд дээр тулгуурлан, NLP системийн чадавхыг сайжруулахын тулд шинэ боломжуудыг эрэлхийлэн судалсаар байна.

2. Судалгааны арга, хэрэглэгдэхүүн

Статистикийн машин орчуулга (SMT) нь зэрэгцээ хөмрөг болох хоёр ба түүнээс дээш хэлээр бичигдсэн, харилцан орчуулагдсан бичвэрийн багц дээр суурилдаг. Эдгээр хөмрөгүүд нь статистикийн загварт зориулсан сургалтын өгөгдөл болно. Зэрэгцээ хөмрөг SMT -ийн суурь болж, харилцан орчуулагдсан олон хэлний эх бичвэрийн түүврийг бүрдүүлдэг. Энэхүү найдвартай өгөгдлүүдийн багц нь хэл хоорондын уялдаа холбоог ойлгох, үнэн зөв орчуулгыг бий болдог юм. Зэрэгцээ хөмрөг бий болгоход нөр их хөдөлмөр ордог бөгөөд янз бүрийн чиглэлээр төрөл бүрийн бичвэрүүдийг нэгтгэн орчуулгыг эмхэтгэх шаардлагатай болдог. Тодорхойлбол зэрэгцээ хөмрөг нь SMT загваруудыг сургахад гол үүрэг гүйцэтгэдэг бөгөөд хос хэлний нарийн ухагдахууныг ойлгож, үнэн зөв орчуулга хийх боломжийг олгодог. Тэдгээрийн хэрэглээ нь машин орчуулга төдийгүй, эх хэлний боловсруулалт, дарааллын загварчлал гэх мэт өргөн хүрээний салбарт хувь нэмэр оруулдаг. Бид математик загварчлал талаас Бэйсийн шийдвэрийн дүрэм болон хамгийн их үнэний хувь бүхий аргыг, компьютерын ухааны үүднээс хиймэл оюуны гол арга болох гүн сургалтын аргыг, гуравдугаарт орчуулгын загварт шаардлагатай математик хэл шинжлэлийн болон хөмрөг хэл шинжлэлийн гол гол аргуудыг сонгон ашигласан. Энэхүү ажил нь өгүүлбэрт тохируулсан эх бичвэр болон зорилтот өгүүлбэрүүдээс бүрдсэн зэрэгцээ хөмрөг шаарддаг хяналттай сургалтын үйл явц. Үүнийг мөн хоёр хэлний хөмрөг гэж нэрлэх ба аливаа эрэмбэлэгдсэн 2 хэлний хувьд бие бие рүүгээ орчуулагдах боломжтой байна гэсэн үг. Хөмрөгийн зорилтот хэл эх өгүүлбэрийн лавлагаа орчуулгыг агуулдаг. Бид системийг сургахад ашигладаг зэрэгцээ хөмрөг сургалтын өгөгдлийн багц (train) болон параметрийг тааруулах эсвэл сургалтын явцад суралцах хурдны хуваарьт өөрчлөлт оруулах зэрэгт ашигладаг жижиг зэрэгцээ хөмрөг хөгжүүлэлтийн өгөгдлийн багц (dev) хоёрыг тус тусад нь бэлтгэж оруулдаг. Гүн сургалт дууссаны дараа орчуулгын системийн гаралтыг үнэлэхийн тулд нэг буюу хэд хэдэн тестийн өгөгдлийн багц (test) ашигладаг. Бид сургалтдаа United Nations Parallel Corpus, Wikimatrix, OpenSubtitles зэрэг нээлттэй эхийн хөмрөгүүдийг орчуулга хийх замаар Англи Монгол хэлний холимог зэрэгцээ

хөмрөгийг бий болгож хэрэглэсэн. Харин яриа таних алхамд бид OpenAI whisper, Coqui, Realtime open source STT загваруудыг хөглөх сургалт хийж оруулсан ба англи хэлээрх танилтын үнэний хувь хэрэглээнд оруулахад хангалттай байлаа.

а. Үгийн оноолт

Үгийн оноолт гэдэг нь эх болон зорилтот дараалал дахь үгсийн хооронд үгийн түвшний хамаарлыг хэлнэ. Ихэвчлэн зэрэгцээ хөмрөгийг үгийн түвшинд харгалзуулан тэмдэглэдэггүй. Тиймээс Үгийн харгалзааг автоматаар тооцдог. Үндсэн санаа нь $T \subseteq \{1, 2, \dots, K\} \times \{1, 2, \dots, L\}$ гэсэн үгийн харгалзаа нь $k \in \{1, 2, \dots, K\}$ эх бичвэрийн индексүүдийн хамаарал гэж тодорхойлдог ба зорилтот өгүүлбэрийн индексүүд $l \in \{1, 2, \dots, L\}$ болно гэсэн үг. Үгийн харгалзааг далд хувьсагчийн дараалал хэлбэрээр (Brown et al., 1993), (Vogel et al., 1996) нэвтрүүлж болно. Энэхүү аргачлалыг ашигласнаар үг, өгүүлбэрийн бүтцийг илүү нарийн тодорхойлж орчуулгын загвартаа тусгах боломж бүрддэг. Бид CoNLL-U (Buchholz & Marsi, 2006) форматыг ашиглан монгол хэлний үг, өгүүлбэрийн бүтцийг тодорхойлсон. Өөрөөр хэлбэл орчуулгын магадлалыг уртын магадлал, харгалзуулах магадлал болон лексик магадлал болгон задалдаг формат гэж ойлгож болно.

б. Орчуулгын загварууд ба хэлний загварууд

Орчуулгын загвар нь тухайн эх хэлний үгс ба тэдгээрийн орчуулгад үндэслэн магадлал бий болдог. Эдгээр магадлал үүсгэхэд сургалтын үе шатанд зэрэгцээ хөмрөг ашиглан сургадаг. Орчуулгын загварууд нь машин орчуулгын салбарт гол үүрэг гүйцэтгэдэг. Сургалтын үе шатанд эдгээр загварууд зэрэгцээ хөмрөг дотроос мэдээлэл цуглуулж, нарийн төвөгтэй хэлний тохиргоог сурч эх болон зорилтот үг, хэллэгүүдийн хоорондын хамаарлын талаарх ойлголтыг гүнзгийрүүлдэг. Орчуулгын загвар нь үндсэндээ хэл шинжлэлийн хяналтын үүрэг гүйцэтгэдэг бөгөөд өгөгдсөн үг хэллэгт өөр өөр орчуулгын магадлалыг тооцно.

Сургалтын үйл явц нь орчуулгын загварын параметруудийг оновчтой болгох, орчуулгын хамгийн боломжит хувилбаруудын талаарх жинд үндэслэн шийдвэр гаргах чадварыг сайжруулахад оршино.

Хэлний загварууд нь зорилтот хэлээрх үгсийн дарааллын магадлалыг тооцоолох үүрэгтэй тул машин орчуулгын салбарын хэсэг юм. Энэхүү үнэлгээ нь зорилтот хэлний дотоод бүтэц, хэв маягт суурилдаг бөгөөд хэлний загварт төрөл бүрийн үгсийн дарааллын магадлалыг тооцоолох боломжийг олгоно.

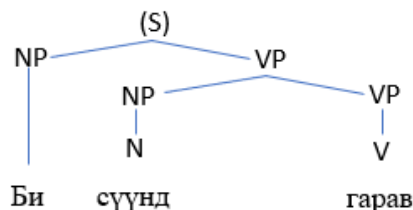
Машин орчуулгын цогц харилцан үйлчлэлд орчуулгын загвар ба хэлний загвар хоорондын ажиллагаа нь динамик буюу үр дүнтэй байдлаар гарч илэрдэг. Өгөгдсөн эх хэллэгийн өөр өөр орчуулгад магадлалыг чадварлаг оноох ба орчуулгын загварыг хэлний загвартай хослуулснаар орчуулга жигд, ойлгомжтой болдог.

Орчуулгын болон хэлний загвараас олж авсан мэдээллийг нэгтгэснээр машин орчуулгын систем нь эх сурвалжийн агуулгыг үнэн зөв гаргаад зогсохгүй зорилтот хэлний онцлог, хэв маягт нийцсэн орчуулгыг бий болгохыг эрмэлздэг. Үүний үр дүнд орчуулга болон хэлний загварчлалын зохицол нь харилцан уялдаатай, нөхцөл байдалд тохирсон бол хэл шинжлэлийн хувьд боловсронгуй орчуулгыг бий болгоно.

с. Монгол хэлний өгүүлбэрийн бүтэц, хүрээ

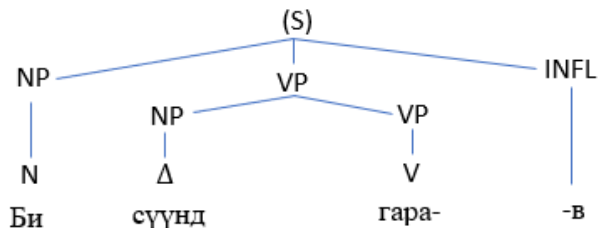
Бичмэл эхийн автомат хэл шинжлэлийн шинжилгээ, тухайлбал синтаксын хамаарлыг задлан шинжлэх, зарим талаараа ярианы хэлний үг зүйн хэсгийг тэмдэглэнэ. Өгүүлбэрийг задлан шинжлэх нь эх хэлний боловсруулалтын (NLP) сонгодог даалгавар бөгөөд хэрэглээний хэл шинжлэлийн салбарт хэлийг цаашид гүнзгийрүүлэн судлах нэг томоохон гарц болдог. Бүрэн хяналттай нэг хэлний өгөгдөлд тулгуурласан задлан шинжилгээнд бид treebank ашиглан CoNLL-

U форматаар үг, өгүүлбэрийн бүтцийг хадгалдаг (Dozat et al., 2017). Өгүүлбэрийн модны хамгийн том бөгөөд нийтлэг хоёр ангилал нь хэлхцийн бүтцийн мод ба хамаарлын мод юм. Хэлхцийн бүтцийн мод (Chomsky, 1956) нь өгүүлбэрээс хэлхэц, хэлхцээс ганц үгтэй болох хүртэл өгүүлбэрийг задлан шинжилдэг. Үг, өгүүлбэрийн бүтцийн хамаарлыг харуулсан мод (Tesnière, 1959) бие даасан үгсийн хоорондын өгүүлбэрийн харилцааг тогтоодог. Өгүүлбэрийг задлан шинжлэхэд өгүүлбэрийн гүн бүдүүвч ашигладаг. Уламжлалт хэл зүйн үүднээс S=“Би сүүнд гарав” гэж өгүүлбэр авбал би гарав, сүүнд гарав гэж задална. Гэтэл би гарав гэсэн нийлэмжийн гарав бүрэлдэхүүн “гара-, -в” гэсэн бие даасан хоёр утгат хэсгээс бүрдсэн гэж үзэх тул -в хэл зүйн бүтээвэр “би-” -тэй хэлбэр хийгээд утгын үүднээс шууд биш дам холбоотой байна. Тэгэхээр түүнийг нийлэмж гэж үзэх нь дутагдалтай. Өөрөөр хэлбэл, түүнийг өгүүлбэрийн бүтэц дэх хамаарах бүтэц гэж үзэх боломжгүй. Харин тухайн өгүүлбэрт цагийн утга илэрхийлсэн -в бүтээвэр нь өгүүлбэрийн бүтцэд чухал үүрэг гүйцэтгэдэг тул залруулах дүрмээр түүнийг нэр, үйлтэй тэнцэхүйц бүрэлдэхүүн хэмээн үзэх болсон. Хэдийгээр Чомскийн өгүүлбэрийн гүн бүтцийн анхны хувилбар нь $S \rightarrow NP VP$ байсан боловч түүнийг хожим дахин залруулан $S = (INFLP)$ гэж томъёолоод залруулах дүрэм хэмээн нэрийдэх болжээ (Равдан, 2020). Эхний томъёоллоор дээрх өгүүлбэр (S)-ийг дүрсэлбэл (Зураг 1):



Зураг 1. Өгүүлбэрийн гүн бүтцийн анхны хувилбар (Равдан, 2020)

Удаах томъёоллоор дээрх өгүүлбэр (S)-ийг дүрсэлбэл (Зураг 2):



Зураг 2. Өгүүлбэрийн гүн бүтцийн залруулах дүрэм (Равдан, 2020)

Монгол хэлний хувьд албан ёсны нийтэд хүртээл болгосон UD (Universal Dependency) хараахан гараагүй байгаа тул энэ загварыг шууд ашиглах хүндрэлтэй байлаа.

d. Нейрон сүлжээний давхарга

Хиймэл оюун ухааны хамгийн чухал ойлголт нь хиймэл нейрон сүлжээ буюу зүгээр л нейрон сүлжээ гэж нэрлэдэг хүний тархинд байдаг механизмуудаас санаа авсан тооцооллын загварууд бөгөөд голчлон ангилах ажлыг загварчлахад ашигладаг. Нейрон сүлжээг 1943 онд анх нэвтрүүлсэн (McCulloch & Pitts, 1943). Эхний жилүүдэд зөвхөн тоглоомын ажилд ашиглах боломжтой байсан ч өндөр уян хатан чанар, хурдан бөгөөд хямд тооцооллын нөөц боломжийн улмаас сүүлийн 20 жилд өргөн хэрэглэгдэх болсон. Эдгээр нь өөр хоорондоо холбогдсон олон перцептонуудаас бүрддэг. Өгөгдсөн перцептронд хангалттай хэмжээний перцептрон холбогдсон бол энэ перцептрон ч бас асах ёстой. Перцептрон бол нейрон сүлжээн дэх хамгийн жижиг нэгж юм. Хүний тархины нейронуудаас санаа авч загварыг нь гаргасан. Энэ загвар бодит нейроноос

бага хүчин чадалтай боловч хиймэл оюуны олон төрлийн хэрэглээнд хангалттай хүчтэй бөгөөд өндөр үр дүнтэйгээр тооцоолох боломжтой (Gidon et al., 2020). Олон давхаргат перцептрон нь нейрон сүлжээний үндсэн архитектур юм (Rumelhart et al., 1986), (Werbos, 1988), (Bishop, 1955). Перцептронууд олон давхаргад хуваагдаж, хоорондоо холбогдсон байдаг. Хэрэв тэдгээрийн мэдээлэл зөвхөн урагшлах боломжийг олгодог давтагдах боломжгүй холбогдсон бол тэдгээрийг дамжуулагч нейрон сүлжээ гэж нэрлэдэг. Хэрэв бид бүрэн холбогдсон сүлжээтэй ажиллаж байгаа бол хоёр давхаргыг холбох нь эхний давхаргын зангилаа бүрийн гаралтыг хоёр дахь давхаргын зангилаа бүрийн оролттой холбохтой тэнцүү юм. Бүрэн холбогдсон сүлжээ нь хамгийн өргөн тархсан сүлжээ юм. Бүрэн холболтууд нь матрицын үржүүлэх хэлбэрээр үр дүнтэй загварчлагдах боломжтой тул тооцооллыг хялбаршуулдаг. Дор хаяж нэг давхаргыг оролтын давхарга болгон ашигладаг. Оролтын давхаргын идэвхжүүлэлтийг тооцооллын оролт болгон вектороор өгдөг. Үүний нэгэн адил загвар нь дор хаяж нэг гаралтын давхарга агуулдаг. Гаралтын давхаргын утгууд нь загварын таамаглал юм. Судалгааг боловсруулах явцдаа нейрон сүлжээний загваруудад ихээхэн анхаарал хандуулсан. Яриа таних, нүүр хувирал таних болон компьютерын алсын хараа зэрэг чиглэлүүдэд хүчтэй нөлөө үзүүлсэн гүн нейрон сүлжээний загваруудыг машины орчуулгын загвар дээр туршиж, хэрэглээнд нэвтрүүлээд удаагүй байна. Нейрон сүлжээг машины орчуулгад ашиглах анхны санааг (Allen, 1987) гаргаж байсныг тэмдэглэх нь зүйтэй. Эхэндээ нейрон сүлжээнд суурилсан загваруудыг үг, өгүүлбэрийн бүтцийг харгалзан үздэг системд нэмэлт загвар болгон нэвтрүүлсэн. Удалгүй нейрон сүлжээнд суурилсан загварууд машины орчуулгын үндсэн хөдөлгүүр болж, энэ аргыг нейрон сүлжээнд суурилсан машины орчуулга гэж нэрлэжээ (Sutskever et al., 2014). Тус арга нь зөвхөн нэг загвар ашиглан машины орчуулгын системийг хялбаршуулдаг учраас бус, харин нэмэлт завсрын алхмуудыг шаардалгүйгээр эцсийн орчуулгыг хурдан шуурхай хийдэг учраас олон нийтэд таалагдсан байж болох. Заавал олон загвар ашигладаг үг, өгүүлбэрийн бүтцийг харгалзан үздэг системтэй харьцуулахад, нейрон сүлжээнд суурилсан машины орчуулга нь илүү энгийн, цэвэрхэн, бүр үг, өгүүлбэрийн бүтцийг харгалзан үздэг системээс ч илүү байдаг. Өнөөдөр судалгааны ажлаа бичиж байгаа яг энэ үед ч нейрон сүлжээнд суурилсан машины орчуулга тооцооллын хэл шинжлэл тэр дундаа машины орчуулгын салбарт давамгайлсан хандлага хэвээр байна.

Гүн нейрон сүлжээнүүд эх хэлний боловсруулалтын даалгавруудад улам бүр түгээмэл болж байна (Akbik et al., 2018). Үгсийг салангид бие даасан хэсэг гэж үздэг тоололд суурилсан загваруудаас ялгаатай. Нейрон сүлжээ нь салангид үгсийг тасралтгүй орон зайд буулгаж, үг бүрийг бодит утгын вектор хэлбэрээр илэрхийлдэг. Эдгээр тасралтгүй дүрслэл нейрон сүлжээг бүхэлд нь сургах процедурын нэг үндсэн хэсэг болно. Яг ийм дүрслэл үг хэллэгийг загварчлахад онцгой ач холбогдолтой байдаг, учир нь үгс тэдгээрийн хоорондын зайг тооцоолох боломжийг олгодог. Үгс тэдгээрийн зайг ашиглан үгс утгын хувьд хоорондоо ойр байгаа эсэхийг тодорхойлох боломжтой. Харьцуулбал, салангид үгст суурилсан загварууд нь үг хоорондын ийм утгын харилцааг тодорхойлж чадахгүй. Үргэлжилсэн орон зайн үгийн дүрслэлийг бий болгоход ашигладаг нейрон сүлжээнээс өөр аргууд байдаг хэдий ч одоогоор орчуулгын салбарт хэрэглээ нь багасаж байгаа (Landauer & Dumais, 1997). Үг хэллэгээс цааш үгийн дараалал руу шилжих үед нейрон сүлжээ нь тоонд суурилсан загварт байдаггүй ерөнхийллийг ашигладаг. Жишээлбэл, тоололд суурилсан хэлний загварууд нь өгөгдлөөс ангид үгсийн дараалалд тэг магадлалын утгыг оноох (Ney et al., 1994) аргыг ашигладаг. Тэг магадлалын утгууд нь олон төрлийн их өгөгдөл, эх бичвэрийг агуулсан статистик өгөгдөлд суурилсан аргууд эсвэл магадлалын гинжин дүрмийг хэрэглэхэд бэрхшээлтэй байдаг тул тэг магадлалаар үржүүлснээр нийт тэг оноо авахад хүргэдэг.

Харин нейрон сүлжээ нь сургалтын явцад ажиглагдаагүй үгийн дарааллын 0-ээс өөр магадлалыг тооцоолох чадвартай. Нейрон сүлжээний загвар нь оролтын утгыг гаралтын утгуудтай харьцуулах функцийг барагцаалагч юм. Энэхүү судалгаанд нейрон сүлжээний оролт нь нэг ба олон үг, өгүүлбэр, эх бичвэр байх бөгөөд гаралт нь сургагдсан загварын хамгийн өндөр магадлалтай дараалал юм. Мэдээж хэрэг бид эх бичвэрийг өгүүлбэрүүдэд задалж, өгүүлбэр бүрийн хувьд нейрон сүлжээний гаралт болгон хувиргаж байгаа.

Нейрон сүлжээ нь ихэвчлэн оролтын, далд, гаралтын давхарга зэрэг олон давхаргат бүтэцтэй байдаг. Бид хэлний загварчлал, машины орчуулгад эдгээр давхаргуудыг нарийвчлан авч үздэг.

3. Дуу, дүрс бичлэг, ярианы англи-монгол машин орчуулга

Орчуулгын загварын магадлалыг A , хэлний загварчлал болон BPE ашигласан илэрцийн загварын магадлалыг B , үг, өгүүлбэрийн бүтэц, өгүүлбэрийн хүрээг тодорхойлох загварын магадлалыг C гэе. Нэгэнт бид монгол хэлний өгөгдсөн өгүүлбэрт хамгийн сайн тохирох англи хэлний өгүүлбэрийг хайж байгаа учир A , B , C -ийн аль алиных нь хувьд хамгийн сайн байх хувилбарыг олох хэрэгтэй.

Одоо байгаа нейрон сүлжээнд суурилсан машины орчуулгын загварууд нь эдгээр гурван загварын нэгдэл байдлаар машины орчуулгын асуудлыг шийдэж ирсэн. Өөрөөр хэлбэл, бие биеэсээ хамаарсан нэг цогц загвар гаргаж авахыг эрмэлзсэн гэж ойлгож болно. Энэ нь нэг талаас машины орчуулгыг судлаач бүр хийх, турших боломжтой болж байгаа ч сургалт явуулах машины хүчин чадал асар өндөр байхыг шаардаж байгаа. Харин бидний хувьд илүү модульчилсан төхөөрөмжийн хүчин чадал бага шаардах загварыг илүүд үзсэн. Учир нь одоогоор машины орчуулгын салбарт монгол хэлтэй хосолсон эрэмбэлэгдсэн хос зэрэгцээ хөмрөг дутмаг, монгол хэлний үг, өгүүлбэр зүйн бүтэц олон улсын нийтлэг UD -д бүртгэгдээгүй, хэлний загварыг BPE ашиглан сургасан туршилт ховор, туршилт явуулах өндөр хүчин чадал бүхий тоног төхөөрөмж дутмаг зэрэгтэй холбоотой. Магадлалын тодорхойлолт ёсоор $P(B/A) = P_A(B)$ нь A нөхцөл дахь B үзэгдлийн магадлал. Одоогоор бидний гаргаж байгаа загвар нь дээр дурдсан 3 загварын шаталсан хэлбэр бөгөөд тус тусдаа үл хамаарах загвараас хамааруулж эцсийн орчуулгаа хийж байгаа юм. Цаашид энэ загварууд дээр өөр нөхцөл нэмэгдэх болгонд тэр болгоны нөхцөлт магадлалыг олох шаардлага гарна. Энэ тохиолдолд бид $A = A_1, B = A_2, C = A_3$ гэх мэтчилэн шаталсан загвар нэмэгдэх тусам нэмэгдсэн тоогоор n хүртэл нөхцөлийг өсгөж болох юм (Томьёо 4).

$$P(A) \cdot P(B/A) = P(B) \cdot P(A/B) \quad (2)$$

Нэгэнт дээрх томьёо хүчинтэй байдаг тул үүнийг дурын $A_1, A_2, \dots, A_n$ -ийн хувьд бодъё.

$$P(A_1 A_2 \dots A_n) = P(A_n / A_1 A_2 \dots A_{n-1}) \cdot P(A_1 A_2 \dots A_{n-1}) \quad (3)$$

Энэ мэтээр $P(A_1)$ хүртэл давтал бидний гаргасан загварын магадлал дараах байдалтай болно.

$$P(A_1 A_2 \dots A_n) = P(A_n / A_1 A_2 \dots A_{n-1}) \cdot P(A_{n-1} / A_1 A_2 \dots A_{n-2}) \cdot \dots \cdot P(A_2 / A_1) \cdot P(A_1) \quad (4)$$

4. Судалгааны үр дүн

Хэлний дүрэм, өгүүлбэрийн хүрээг загварчлахдаа эхлээд монгол хэлний өгүүлбэрийн ерөнхий хамаарлыг графикаар дүрсэлсэн. Түүний дараа CoNLL-U форматаар хадгалж Stanford загвар ашиглан сургаж боловсруулалтдаа ашигласан. Зарим тохиолдолд олон хэлээр ярьж байгаа тохиолдолд хэлийг таних шаардлагатай тулгарсан ба language detector программыг орчуулгын программдаа мөн нэмсэн. Бид судалгаандаа дараах Python хэлний сангуудыг ашигласан.

```

from pydub import AudioSegment
import math
import requests
import os
import json
from textwrap import wrap
from pvrecorder import PvRecorder
import wave
import struct

Яриаг орчуулах туршилтыг хийхдээ микрофон ашиглан аудиог бичиж авсан.
recorder = PvRecorder(device_index=-1, frame_length=512)
audio = []
try:
    recorder.start()
    while True:
        frame = recorder.read()
        audio.extend(frame)
except KeyboardInterrupt:
    recorder.stop()
    with wave.open(path, 'w') as f:
        f.setparams((1, 2, 16000, 512, "NONE", "NONE"))
        f.writeframes(struct.pack("h" * len(audio), *audio))
finally:
    recorder.delete()
sound = AudioSegment.from_wav('file.wav')
sound.export('file.mp3', format='mp3')
split_audio("file.mp3")
Дуу хоолойг дараах кодыг ашиглан хуваасан.
ig = 0
# Path
def split_audio(file_path, segment_length=0.5*60*1000): # 30 seconds in milliseconds
    audio = AudioSegment.from_file(file_path)
    total_length = len(audio)
    num_segments = math.ceil(total_length / segment_length)
    for i in range(num_segments):
        start_time = i * segment_length
        end_time = min((i + 1) * segment_length, total_length)
        segment = audio[start_time:end_time]
        output_file = f'{file_path[:-4]}_part{i+1}.mp3'
        segment.export(output_file, format="mp3")
        print(f"Exported: {output_file}")
        print(f"i=",i)
    global ig
    ig = i

```

Орчуулгын программыг бид боловсруулж дотоод сүлжээнд байршуулсан ба Whisper API, UNTranslate API, Language Detector зэргийг холбож ашигласан. Энэ нь хөгжүүлэлтийн хурдыг эрс нэмэгдүүлсэн ба серверийн хүчин чадлыг зөв хуваарилахад тус дөхөм болсон. Туршилтын серверийн хүчин чадлаас хамааран ажиллах хурд бусад системүүдтэй харьцуулахад удаан байсныг энд тэмдэглэе.

5. Дүгнэлт

Дуу, дүрс бичлэг, ярианы зөв хуваарьтай хадмал тайлбар үүсгэж чадвал зөв орчуулсан хадмал орчуулгыг автоматаар үүсгэж болно. Өөрөөр хэлбэл өгүүлбэрийн төгсгөлийг зөв таньж, хадмал тайлбарыг зөв үүсгэснээр орчуулгын чанарт эергээр нөлөөлж байгаа нь туршилтын үр дүнгээр харагдлаа. Ялангуяа өргөн хэрэглээний бичвэр орчуулахад нэн тохиромжтой нь судалгааны үр дүнгээр батлагдсан. Нөгөө талаас олон хэлээрх яриаг таньж зөв орчуулах үед эхлээд хэлийг маш зөв таних шаардлагатай тулгарч байлаа. Цаашид энэхүү туршилт судалгааны үр дүнг ашиглан олон хэлний яриаг автоматаар таньж орчуулдаг машин орчуулгын систем боловсруулахаар ажиллаж байна.

Ном зүй

- Равдан, Э.** (2020a). *Хэрэглээний хэл шинжлэл*. Улаанбаатар: Удам соёл хэвлэх үйлдвэр. [Ravdan, E. (2020). *Kheregleenii khel shinjlel*. [Applied Linguistics] Ulaanbaatar: Udam soyol LLC.]
- Akbik, A., Blythe, D., & Vollgraf, R.** (2018). Contextual String Embeddings for Sequence Labeling. *27th International Conference on Computational Linguistics*, 1638–1649. Retrieved from <https://github.com/zalandoresearch/flair>
- Allen, R. B.** (1987). Several Studies on Natural Language and Back-Propagation. *IEEE International Conference on Neural Networks*, 335–341.
- Aman Priyanshu**, Yash Maurya, Zuofei Hong. (2023). AI Governance and Accountability: An Analysis of Anthropic’s Claude
- Alec Radford**, Jong Wook Kim, Tao Xu, Greg Brockman, Christine McLeavey, Ilya Sutskever. (2022). Robust Speech Recognition via Large-Scale Weak Supervision. arXiv
- Bahdanau et al.** Neural Machine Translation by Jointly Learning to Align and Translate 2015
- Batsukh, BE.** (2023). Sentence Structure and Boundary for Deep Neural Machine Translation Alignment Model. In: Arai, K. (eds) *Proceedings of the Future Technologies Conference (FTC) 2022*, Volume 3. FTC 2022 2022. Lecture Notes in Networks and Systems, vol 561. Springer, Cham. https://doi.org/10.1007/978-3-031-18344-7_36
- Bishop, C. M.** (1955). *Neural Networks for Pattern Recognition*. CLARENDON PRESS. OXFORD.
- Brown, P. F., della Pietra, S. A., della Pietra, V. J., & Mercer, R. L.** (1993). The Mathematics of Statistical Machine Translation: Parameter Estimation. *Computational Linguistics*, 19(2), 263–311. Retrieved from <https://aclanthology.org/J93-2003>
- Buchholz, S., & Marsi, E.** (2006). CoNLL-X Shared Task on Multilingual Dependency Parsing. *Proceedings of the Tenth Conference on Computational Natural Language Learning (CoNLL-X)*, 149–164. Retrieved from <https://aclanthology.org/W06-2920>
- Chomsky, N.** (1956). Three models for the description of language. *IRE Transaction on Information Theory*, 11–124.
- DeepSeek-AI**, (2025). DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning

- Dozat, T., Qi, P., & Manning, C. D.** (2017). Stanford's Graph-based Neural Dependency Parser at the CoNLL 2017 Shared Task. *Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, 20–30. <https://doi.org/10.18653/v1/K17-3002>
- Gemini team and Google.** (2024). Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context
- Gidon, A., Zolnik, T. A., Fidzinski, P., Bolduan, F., Papoutsis, A., Poirazi, P., Holtkamp, M., Vida, I., & Larkum, M. E.** (2020). Dendritic action potentials and computation in human layer 2/3 cortical neurons. *Science (New York, N.Y.)*, 367(6473), 83–87. <https://doi.org/10.1126/science.aax6239>
- Hideyuki Tachibana,** Katsuya Uenoyama, Shunsuke Aihara, (2017). Efficiently Trainable Text-to-Speech System Based on Deep Convolutional Networks with Guided Attention. arXiv
- Landauer, T. K., & Dumais, S. T.** (1997). Psychological Review A Solution to Plato's Problem: The Latent Semantic Analysis Theory of Acquisition, Induction, and Representation of Knowledge. *Psychological Association, Inc*, 104(2), 211–240.
- Llama Team and Meta.** (2023). The Llama 3 Herd of Models
- Mcculloch, W. S., & Pitts, W.** (1943). A LOGICAL CALCULUS OF THE IDEAS IMMANENT IN NERVOUS ACTIVITY. *BULLETIN OF MATHEMATICAL BIOPHYSICS*, 5.
- Ney, H., Essen, U., & Kneser, R.** (1994). On structuring probabilistic dependences in stochastic language modelling. *Computer Speech & Language*, 8(1), 1–38. <https://doi.org/https://doi.org/10.1006/csla.1994.1001>
- OpenAI.** (2023). GPT-4 Technical Report
- Peters et al.** Deep Contextualized Word Representations. 2018.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J.** (1986). Learning representations by back-propagating errors. *Nature*, 323, 533–536.
- Sutskever, I., Vinyals, O., & Le, Q. v.** (2014). Sequence to Sequence Learning with Neural Networks. *CoRR*, abs/1409.3215. Retrieved from <http://arxiv.org/abs/1409.3215>
- Tesnière, L.** (1959). *Eléments de syntaxe structurale* 'Editions Klincksieck. In *Canadian Journal of Linguistics/Revue canadienne de linguistique* (Vol. 6, Issue 1). Cambridge University Press. <https://doi.org/10.1017/S0008413100018922>
- Vogel, S., Ney, H., & Tillmann, C.** (1996). HMM-Based Word Alignment in Statistical Translation. *International Conference on Computational Linguistics*, 836–841.
- Vaswani et al.** Attention is All You Need. 2017.
- Wang et al.** Learning Deep Transformer Models for Machine Translation 2019.
- Werbos, P. J.** (1988). Generalization of backpropagation with application to a recurrent gas market model. *Neural Networks*, 1(4), 339–356. [https://doi.org/https://doi.org/10.1016/0893-6080\(88\)90007-X](https://doi.org/https://doi.org/10.1016/0893-6080(88)90007-X)