

Компьютерын ухаан

Уртын дууны төрлийг шүлгийн өгөгдөл ашиглан машин сургалтын аргаар ангилах нь

Н.Баярсайхан¹, Б.Сувдаа^{1,*}, Г.Нандинбилэг²

¹МУИС-ХШУИС, Мэдээлэл, Компьютерын Ухааны Тэнхим, Машин оюуны лаборатори

²МУИС-ШУС-ХУС, Утга зохиол, урлаг судлалын тэнхим

Received on 2022.04.04; Revised on 2022.05.21; Accepted on 2022.11.03

*Corresponding author: suvdaa@seas.num.edu.mn.

Хураангуй

Түүх, шастираас үзвэл Монгол ардын уртын дуу нь нэн эртний улбаатай аман зохиолын томоохон төрөл зүйлийн нэг юм. Арвин сан хөмрөгтэй Монгол ардын аман зохиол нь хэдэн өдрөөр хайлдаг хөлгөн их туйлсаас эхлээд ерөөл, магтаал, хэдэн арван шад шүлэг бүхий уртын дууны шүлэг зэрэг нь амнаас ам дамжин өнөө бидэнд уламжлагдан ирсэн бөгөөд бүтнээр нь өгүүлж байж төгөлдөр утга бүрнээ илэрхийлэгддэг. Уртын дуугаа гадаад ертөнцөд танин мэдүүлэх нь зөвхөн аялгуу, эгшгийн хүрээнд л байгаа юм. Тэгвэл шүлгийн утга, агуулгаас, ёс заншил, Монголчуудын оюун сэтгэлгээг тайлж мэдэх учиртай юм. Үүнийг орчин цагийн мэдээллийн технологийн хөгжил, судалгааны аргыг ашиглан шинэлэг аргаар судлан, хөгжүүлэх боломжийг эрэлхийлэн энэхүү судалгааг салбар дундын хэлбэрээр эхлүүлж байна. Энэхүү өгүүлэлдээ уртын дууны хэмжээг илэрхийлдэг айзам-12, суман-64, бэсрэг-7 дууг дууны нэр, утга, шүлэг, тухайн дууны бадгийн тоо, мөрийн тоо, үгийн тоо зэргээр бүртгэлжүүлэн машин сургалтын аргаар ангилах, мөн аялгууны үргэлжлэх хугацааны хувьд бадгийн үргэлжлэх хугацаа, 5 үгийн үргэлжлэх хугацаа, 1 үгийн үргэлжлэх дундаж хугацаа, хамгийн урт дуулагдаж байгаа үгийн үргэлжлэх хугацаа зэрэг үзүүлэлтээр статистик дүгнэлт хийлээ. Туршилтын үр дүнгээр Weka програм дээр машин сургалтын аргууд зөвхөн шүлгийг өгөгдөлд дундажаар 60-70 хувь, хамгийн сайн нь ZeroR арга 77 хувь зөв танилттай байгаа нь боломжийн эхлэл гэж дүгнэж байна. Харин аялгууны өгөгдлөөр бол маш амархан ялгарахаар харагдаж байна.

Түлхүүр үг: уртын дууны төрөл ангилал, аман зохиол, монгол шүлэг, машин сургалт

1 Удиртгал

Монголчуудын аман зохиолын төрлөөс тэдний аж байдал, түүх соёлыг харж болохоор байгаа нь уртын дууны яруу найргийг уншихад мэдэгддэг. Уртын дууны шүлгийг анх хэн нэгэн хүн зохиож, улам хөгжин, ам дамжсаар ардын дунд түгэн дэлгэрсэн аман зохиолын шинжтэй болж ирсэн. Монголчуудын бичиг соёлын түүхэн улбааг харахад нэн эртний сурвалжууд бий агаад уртын дууны шүлэг мөн ялгаагүй эртнээс гаралтай, өвөг дээдсийн зохион шүлэглэсэн яруу тансаг шүлэг, яруу найргийн дээж болно. Тэгвэл өнөө цагт уртын дууг бүх түрлэгээр дуулдаггүйгээс уртын дууны яруу төгс утга бүрнээ сонсогдохгүй, мэдэгдэхгүй өнгөрсөөр байна. Бид энэ судалгаагаар уртын дууны яруу найргийн утгыг илэрхийлж утга зохиолын талаас тайлбарлахаас гадна уртын дууны үгийн өгөгдлийг машин сургалтын арга ашиглан айзам, суман, бэсрэг гэсэн 3 ангилалд ялган таниулж, дүгнэлт хийв. Уртын дууны яруу найраг нь бичгийн шинж давамгайлсан, яруу сэтгэлгээний дуулал маягтай байдаг бөгөөд нэн ялангуяа айзам уртын дуунд ертөнц дэл-

хийн үзэгдэл, хүний амьдралын жам ёсны тухай гүн ухааны шинжийг төгс агуулсан байдаг бол суман уртын дуунд эгэл ардын аж төрөх ёс, ахас дээдэс, ижий, ааваа хүндлэх, уул, амраг сэтгэлээ илэрхийлэх, усаа дээдлэх ухаан, мал, аж ахуй зэрэг түгээмэл сэдвээс гадна Монголчуудын түүхийн ороо, бусгаа цагийг ч танин харж болохоор өргөн дэлгэр агуулгатай байна. Харин бэсрэг дуунд захин, сургасан шинж давамгайлж байдаг зэргээс харахад Монголчуудын аль ч үеийн түүхийг уртын дууны шүлэг, яруу найргаас харж болох юм. Машин сургалтын аргаар дууны шүлгийг ангилан танин чиглэлийн судалгаа дэлхий нийтэд нэлээд хийгдэж байгаа нь олон салбарт ажиглагдаж байгаа бөгөөд Монгол хэл шинжлэлийн салбар нь цахим хэл шинжлэл талд салбар дундын судалгааг түлхүү хөгжүүлэн, чамгүй ололт, амжилт гаргаад байна. Тэгвэл аман зохиол, утга зохиолын салбарт салбар дундын ийм судалгаа анхдагч болж байгаагаараа бидний энэхүү ажил шинэлэг юм. Тэр дундаа уртын дууны шүлэг, яруу найргаар эхлүүлж байгаа нь бэлгэдлийн хувьд ч ач холбогдолтой. Тиймээс орчин цагт дэлхий нийтэд олон талын судалгаанд өргөнөөр ашиглагдаж буй

хиймэл оюун ухааныг ардын аман зохиолын үнэт өвтэй хослуулан судалж, хойч үедээ сурталчлан таниулах нь бидний нэгэн зорилго юм. [1-3]

2 Судлагдсан байдал

Уртын дууг шинжлэх ухааны үндэслэлтэйгээр анх судалсан хүн бол Оросын эрдэмтэн А.Позднеев юм. Тэрээр 1880 онд “Монгол уртын дуу” бүтээлдээ буриад, халх, өөлд дуунуудыг багтаасан байдаг бол мөн Оросын эрдэмтэн А.Д.Руднеев монгол уртын дуунуудын ая эгшгийг судалсан бол Б. Я. Владимирцов Ойрад дууг нарийн судалж тэмдэглэсэн байна. [1] Уртын дууны аман эхээс гадна бичмэл эх энэ салбарын судалгааны голлох хэрэглэгдэхүүн болдог. П.Хорлоо, Х.Сампилдэндэв [4], Ш.Дорж, Ж.Бадраа, С.Цоодол нарын зэрэг Монгол эрдэмтэн, судлаачид уртын дууны талаар судалсан бөгөөд Монгол нутгаар нэлээд тархмал байсан бичмэл судар, номуудыг ШУА-ийн ХЗХ-ийн зарим ахмад эрдэмтэд цуглуулсан байдаг. Мөн алдарт уртын дуучин Ж.Дорждагва [5] уртын дуу судлалын түүхэнд томоохон орон зай эзлэх судалгааг хийсэн. Ж.Бадраа түүнтэй хүүрнэсэн яриаг “Их дуучны яриа” нэртэй ном болгон хэвлүүлсэн нь энэ салбарын эрдэмтэн, судлаачдын дунд үнэ цэнтэй бүтээл болдог. Мөн судлаач А.Алимаа [1] Монгол нутагт дуулагдаж байсан нийт 3000 гаруй уртын дууг судалж, нээн илрүүлж, судалгааны эргэлтэд оруулсан нь судлаачдын хувьд чухал хэрэглэгдэхүүн болсон. Тэгвэл Монгол шүлэг судлалын хувьд XIX зууны дунд үеэс эхлэн өрнөдийн Монгол судлаачид судалж эхэлсэн түүхтэй. Тэд Монгол шүлгийн онцлог толгой холбоцод нь байгааг ажиглан, онцлон өгүүлсээр иржээ. Тухайлбал Оросын Монгол судлаач А.Позднеев “Монгол шүлгийн гол онцлог толгой холбох явдал”, мөн Оросын Монгол судлаач Н.Поппе “Монгол туулийн бадагт толгой холбосон аллитерацаас гадна дотоод аллитераци ч зарим үед харагдаж байна.”, Чехийн Монгол судлаач П.Поуха “Монгол шүлгийн гадаад гол шинж нь гэвэл толгой холбож, ая нийлүүлэх явдал юм. Ийм шинж Монгол шүлэглэлд эрт дээр үеэс тодорчээ” гэх мэтээр судалгааны бүтээлүүдэд дурдсанаас харж болно. [2] Олон улсад дууны төрлийг ангилал (genre classification) судалгааны ажлууд нэлээд өргөн хийгдэж, өнөө цагт хэрэглээнд нэвтрээд эхэлсэн байна. Ингэхдээ дууны аудионы өгөгдөлд дохио боловсруулалт хийн рок, поп, реп, сонгодог гэх мэт төрлүүдийг ангилсан ажлууд түгээмэл бий. [6-10] Харин зөвхөн шүлгэнд тулгуурлан ангилал хийсэн ажил аудионоос цөөн боловч, эх хэл боловсруулалт, машин сургалт ашиглан хийсэн ажлууд бас байна. Сүүлийн жилүүдэд гүн сургалтын хэрэглээ ихэссэн бөгөөд үүнийг дагаад шүлгийн өгөгдөлд RNN гэх мэт гүн сургалтын аргаар төрлийн ангилал хийж байна. Тухайлбал Alexandros Tsapsinos нарын ажил [6] харж болох юм. Мөн Anna Boonyanit нарын ажил [7] хип-

хоп, рок, поп төрлүүдийг 60 орчим хувь зөв танилгатай ангилсан байна. Харин уртын дууны агуулгын ангилал хийгээд шүлгийн утгыг цахимаар таниулж үзсэн судалгаа бидний судалснаар одоогоор Монголд ховор байна. [1,11]

3 Уртын дууны шүлэг зүй

Монгол ардын уртын дуу нь Монгол аман зохиолын томоохон төрөл зүйлийн нэг юм. Аман зохиол нь ард олны аж амьдралаас үүдэн зохиогдож амнаас ам дамжин Монгол хүний зан заншил, түүх соёл, мэргэн ухаан, цэцэн билгийн илэрхийлэл болж уламжлан дэлгэрсэн уран үгийн бүтээл юм. Ардын аман зохиолын үндсэн төрөлд үлгэр, тууль, домог, оньсого, таавар, зүйр цэцэн үг, ерөөл, магтаал, ертөнцийн гурав, хууч яриа зэрэг багтана. Эдгээр төрөл зүйлүүдийн ихэнх нь шүлэг, яруу найргийн шинжтэй байдаг. Хүний үүслийн өлгий нутаг гэгддэг Монгол нутагт яруу найргийн баян тансаг өв, уламжлал бий. Хэдхэн үгтэй дууны шүлгээс эхлэн олон хоног аялгуулан хайлдаг туульс хүртэл Монгол шүлэг, яруу найргийн баян тансгийг харах боломж бий. “Монголын нууц товчоо”, “Жангар”, “Гэсэр” зэрэг хөлгөн их туульсаас гадна Чингис хааны “Оюун түлхүүр” сургааль, “Чингисийн есөн өрлөгтэй өнчин хүүгийн цэцэлсэн шастир”, “Эр хоёр загалын тууж” Тогоонтөмөр хааны “Гэмшлийн шүлэг”, Сөнөдийн Гилгүдэй баатрын “Харууслын шүлэг”, Мандухай цэцэн хатны “Тангарагийн шүлэг”, Ноён хутагт Равжаагийн шүлгүүд зэрэг олон арван шүлэг, яруу найраг, тууль Монголын яруу найргийн үнэт өв болон өдгөө бидний үед уламжилжээ. Энэ мэтээр Монгол шүлэг, яруу найргийн өв уламжлал нэн арвин агаад үүн лугаа уртын дууны шүлгийн утга агуулга, зан заншил холбогдоно. Уртын дууны яруу найраг нь бичгийн хэлний үг найруулгаар голдуу бүтсэн байгаа нь анхаарал татах зүйл мөн. Монгол бичгийн эрхэмсэг найруулга, монгол хэлний сангийн ховор нандин үгсийг сонгосон байдал нь монгол уртын дуу нь аман зохиолын төрөл зүйл төдийгүй бичгийн яруу найргийн шинжтэй болохыг харуулж байна. Уртын дууны яруу найргийн олонхийг бичгийн өндөр боловсолтой хүмүүс зохиосон гэлтэй. Ялангуяа төрт ёсны холбоотой хүндэтгэлийн уртын дуунд энэ нь улам ч тодорхой ажиглагдана. Орчлон ертөнцийн юмс үзэгдэл, байгалийн жам ёс, зан заншил, хүндлэл хайр гэх мэт олон талын мэдлэг, сургаалийг олж авахаар төгөлдөр утга агуулга уртын дууны шүлгээс уншигддаг нь дээрхийн илэрхийлэл юм. [12, 13] Чингис хаан болоод Монголын эзэнт гүрний их хаадаас төрийн албан захиа бичиг сэлтийг илгээхдээ хүргэх элчид цээжлүүдэд явуулдаг байсан гэдэг. Нэг талаар захиа бичиг үрэгдэх осолдох магадлалыг тооцсон байх боловч гол нь өөр үүрэгтэй байв. Элч уран үгээр шүлэглэсэн захиаг хүргэж очоод удирдагч, жанжин хэнд ч болов дуудан хэлж өгдөг байж. Энэ нь уг бичиг

захидлын утга агуулгыг бүрнээ хүргэхийн хамт түүгээр дамжуулж буй сэтгэлийн хөдлөл, тухайн цаг төрийн нөхцөл байдал зэрэг олон хүчин зүйлийг хүргэх бололцоотой бөгөөд үгийн хүчийг улам эрчимжүүлэх арга билгийг монголчууд сайн мэддэг байсантай холбоотой. Энэ байдал уртын дууны утга агуулга, шүлэгт гүн тусгагдсан байдаг. Уртын дууны шүлгийг судлахад аман зохиолын онолоос гадна Монгол шүлгийн зүй тогтлын тухай судлах хэрэгтэй юм. Монгол шүлэг нь хос уянга шинжтэй бөгөөд аливаа нэг зүйлийг нөгөөтэй хослуулсан, харьцуулсан шинж давамгайл байдаг. Мөн толгой, дунд холбоц, сүүл холбоц түгээмэл ажиглагддаг нь аман зохиолын дээрх төрлүүдэд бүгдэд нь бий. Монгол шүлэгт хос уянгыг нэгэн цогц утгад зангидан дахин ижил хос уянгатай ижилсүүлэн өрөх маягаар дөрвөн мөр шүлгийг үүсгэдэг зарчимтай. Ийм шүлгийн сонгодог жишээ бол ардын дуу юм. Шүлгийн холбоц нь өнгөн талдаа шүлгийн хэлбэрийг төгөлдөржүүлэх мэт боловч дотоод мөн чанартаа туурвиж байгаа үндэстнийхээ сэтгэлгээний уугуул хэв загварыг илэрхийлдэг тул Монгол шүлгийн толгой холбоц ёс нь Монгол хүний аливаа зүйлийн эхлэлийг эрхэмлэн чухалчилж үздэг үзэлд суурилсан байдаг. Шүлгийн мөрүүдийн эхлэлийг ижил, ойролцоо, эсрэг дуудлага, утгатай үе, үгээр холбож шүлгийн толгой холбоцыг хийдэг. Монголын аль ч цаг үеийн аман хийгээд бичгийн уран зохиол дахь шүлгийг авч үзэхэд ихэнх нь толгой холбогдож бичигдсэн байна. [2] Монголчууд аливаа зүйлд бэлэгшээж ханддаг энэ чанар уртын дууны аялгуунд ч нэвт шингэсэн нь бий. Тухайлбал зарим дууны эхлэл хэсгийн аялгуу ижил байдаг бол заримын төгсгөл ижил дуулагдах тохиолдлууд бий бөгөөд дундаа ижил аялгуутай уртын дуу гэж байхгүй. Жишээлбэл: Дөрвөн цаг, Өвгөн шувуу хоёр дууны эхлэл, Сэрүүн сайхан хангай, Уяхан замба тивийн наран дууны эхлэл гэх мэтээр судалбал олон дууны эхлэл, төгсгөлүүд ижил байдаг. Монгол хүний аливаа зүйлийг өөр бусад ямар нэгэн зүйлтэй харьцуулан жишиж дундаас нь мөн чанарыг нь ухааж ойлгох сэтгэлгээтэй холбоотойгоор хос уянгын аль нэг мөр (ихэвчлэн эхний мөр) ертөнцийн мөн чанарын тухай байх бол нөгөө нь хүний амьдралын утга учрын тухай өгүүлдэг тул харьцуулан жишиж буй хоёр зүйлийг шүлэглэхдээ толгой холбоцын аргаар (хөл холбоц, дунд холбоц ч түгээмэл бий) ижилсүүлэн холбож нэг цогц утгыг илэрхийлдэг байна. [2] Тухайлбал: “Оломгүй далай” уртын дууны төв халх хувилбарт: Оломгүй далай нь Онгоцоо дүүрээд мэлтэрнэ Орчлон дэлхийн хүмүүн нь Мөнхөрнө гэдэг худлаа шүү Үнэгэн зоотой морийг Олж унах нь бэрх хөө Үнэн сэтгэлтэй хүмүүнийг Олж ханилах нь бэрх хөө Туулай зоотой морийг Олж унах нь бэрх хөө Туйлбар сайтай хүмүүнийг Олж ханилах нь бэрх хөө гэсэн шүлгийн эхний бадагт хүмүүний амьдралын мөнх бусын тухай өгүүлэхдээ шүлгийн толгой холбох ёсыг баримталсан байна. Дараагийн бадгуудад сайн, сайхан зүй-

лийг олж хүртэх амаргүй байдаг тухай утга агуулгыг илэрхийлэхдээ толгой холбох болон хөл холбох аргуудыг ашиглажээ. Эндээс ертөнцийн мөн чанар, хүмүүний амьдралын утга учрын талаар харьцуулан жишсэн сонгодог жишээг харж болно. Энэ нь Монгол шүлгийн толгой холбох зарчим ямар ч үед өөрлөгддөггүй, тогтвортой, бат бөх байж ирснийг давхар харуулна. “Оломгүй далай” хэмээх энэ дуу нь хэмжээний хувьд суман дуу юм. Уртын дууг аялгуу, хэмжээгээр нь айзам, суман, бэсрэг гэж гурав ангилдаг бөгөөд шүлгийн утга, агуулгад энэ ангилал мөн л илэрч байдаг. Монгол дууны яруу найргийг судалсан олонх бүтээлд зөвхөн шүлгийн утга агуулгын үүднээс сэдэвчлэн ангилсан байдаг байна. [4, 5, 14, 15] Айзам (том хэмжээний дуу) уртын дууны аялгуу нь цараа өргөн, аялгуу уужим, давхар шуранхай зэрэг чимэг нугалаа ихтэйгээрээ бусад хоёр ангиллаас том хэмжээтэй байдаг бөгөөд үгийн утга агуулгын хувьд ертөнц дэлхийн үзэгдэл, хүмүүний амьдралын жам ёсны тухай гүн ухааны шинжийг төгс агуулсан байдаг. Суман (дундаж хэмжээний дуу) уртын дуу нь уужим тэнүүн аялгуутай, шуранхай, чимэг ихтэй, Монгол оронд өргөн дэлгэр дуулагддаг бол үгийн утга агуулгын хувьд эгэл ардын аж төрөх ёс, ахас дээдэс, ижий, ааваа хүндлэх, амраг сэтгэлээ илэрхийлэх, уул, усаа дээдлэх ухаан, мал, аж ахуй зэрэг түгээмэл сэдвээс гадна Монголчуудын түүхийн ороо, бусгаа цагийг ч танин харж болохоор өргөн дэлгэр агуулгатай байна. Бэсрэг (бэсрэг буюу бага хэмжээний дуу) уртын дуу нь ардын богино дууг бодвол уужим аялгуутай ч шуранхайгүй, богино хэмжээгээр цохилог, чимэглэлтэй дуулагддаг бол үгийн утга, агуулгын хувьд захин, сургасан шинж давамгайлж байдаг. Ийм төрлийн дууг уртын дуу сурч эхэлж буй хүнд сургалтын хэрэглэгдэхүүн болгон ашиглах уламжлал бий. Эдгээр дууг сонсоод аялгуугаар нь дээрх 3 ангиллаар салгаж ойлгож болохоос гадна дууны амин чухал хэсэг болох үгийн утгаар нь ангилах боломж бас бий юм. Тэгвэл үүнийг цахимаар ангилж үзэх нь бидний дараагийн ажил болоод байна. [1, 16]

4 Дууны шүлгийн өгөгдлөөр төрлийг ялган ангилах боломж

Монгол ардын уртын дуу 2008 онд Юнеско-д бүртгэгдэн дэлхийд данслагдсан. [17] Энэ хүртэл аман зохиолын өв талаас нь гадна, дотны чамгүй судалгаа хийгдсэн. Тэгвэл уртын дууг дэлхий дахинд таниулан судалж, судалгааны орчин үеийн арга технологитой хослуулж, салбар дундын судалгаа болгон өргөжүүлэх нь шинэлэг агаад чухал ач холбогдолтой болоод байна. Энэ талын судалгаа хийхэд хамгийн чухал нь ихээхэн хэмжээний өгөгдөл цуглуулсан байх ёстой байдаг. Тэгвэл уртын дууны бичмэл эхүүдийн хувьд нэлээд их судлан цуглуулсан

өгөгдөл байдаг бөгөөд бид энэхүү судалгаандаа төв халх дууны жишээн дээр туршилт хийхээр шийдсэн. Халх уртын дуу нь Монгол орны цөм хэсэгт ихээр тархсан байдаг болохоор одоо цагт нийтлэг дуулагдаж буй ихэнх дуу энэ ангилалд багтдаг. Бид энэ удаагийн ажлаар уртын дууны шүлгийн өгөгдөлд тулгуурлан суман, айзам болон бэсрэг гэсэн 3 төрлийг машин сургалтын аргаар ангилах боломжийг судаллаа. Дараах зурагт ажлын ерөнхий бүдүүвчийг харуулав.

5 Үр дүн

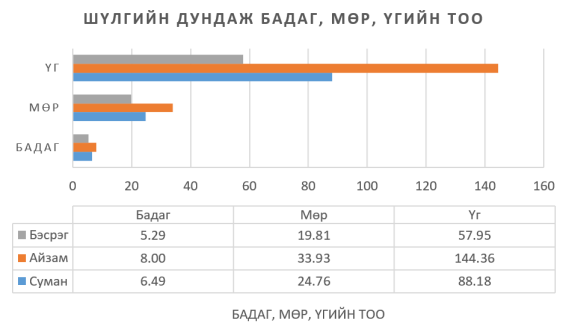


Зураг 1: Уртын дууны шүлгийн төрөл ангилах бүдүүвч

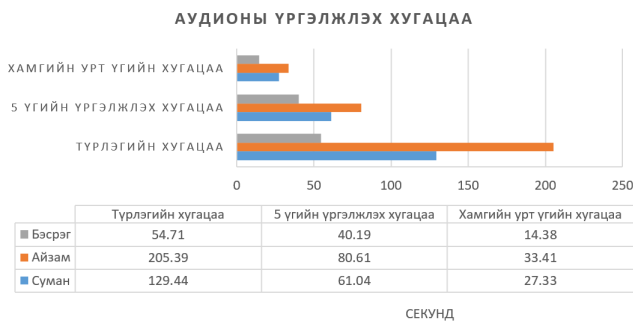
Өгөгдөл бэлтгэх ба онцлог шинж: Бид энэ удаа суман төрлийн 64, айзам төрлийн 12, бэсрэг төрлийн 7, нийт 83 уртын дууны үгийг хэвлэмэл эх сурвалжуудаас нэг бүрчлэн MS Excel програм дээр шивж оруулан, дууны ангилал, утга зэргийг заримыг ном эх сурвалжаас авсан, эх сурвалжид байхгүй заримыг судлаач өөрөө шинжлэн ямар төрөл болохыг тодорхойлж оруулан сургалт, туршилтын өгөгдөл болгон бэлтгэлээ. Цаашид түгээмэл дуулагдаж буй 300 гаруй дууны үгэн дээр энэ туршилтаа хийх боломж бий юм. Бидний цуглуулсан шүлгийн мэдээллээс хамгийн олон давтагдсан тоог гарган үзвэл сайхан, хоёр, морь гэх мэт үгс байна. Зураг 2-т харуулсанчлан хамгийн олон давтагдсан үг хамгийн томоор харагдаж байна. Мөн уртын дууны шүлгийн өгөгдлөөс дээрх 3 ангилалд харгалзах бадаг, мөр, үгийн тоогоор дундажлан Зураг 3-т харуулав. Энд суман дууны бадгийн дундаж тоо 6-7, мөрийн тоо 24-25, үгийн тоо 88-89, Айзам дууны бадгийн дундаж тоо 8, мөрийн тоо 33-34, үгийн тоо 144-145 байгаа бол бэсрэг дууны бадгийн дундаж тоо 5-6, мөрийн тоо 19-20, үгийн тоо 57-58 байгаа нь дээрх графикаас харагдаж байна. Тэгвэл уртын дууны урт гэдэг утгыг илэрхийлэх гол үзүүлэлт болох аудионы хувьд дараах Зураг 4-т харуулав. Энд Суман дууны бадгийн дундаж үргэлжлэх хугацаа 129-130 секунд, 5 үгийн үргэлжлэх хугацаа 60-61 секунд, хамгийн



Зураг 2: Уртын дууны шүлгийн үгийн давтамж



Зураг 3: Уртын дууны шүлгийн өгөгдөл 3 ангиллаар



Зураг 4: Уртын дууны дууны аудионы өгөгдөл 3 ангиллаар

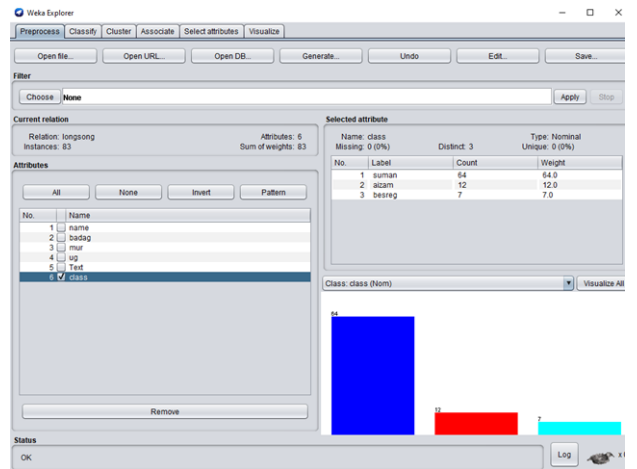
урт үргэлжлэх 1 үгийн хугацаа 27-28 секунд, Айзам дууны бадгийн дундаж үргэлжлэх хугацаа 204-205 секунд, 5 үгийн үргэлжлэх хугацаа 80-81 секунд, хамгийн урт үргэлжлэх 1 үгийн хугацаа 33-34 секунд, байгаа бол бэсрэг дууны бадгийн бадгийн дундаж үргэлжлэх хугацаа 54-55 секунд, 5 үгийн үргэлжлэх хугацаа 40-41 секунд, хамгийн урт үргэлжлэх 1 үгийн хугацаа 14-15 секунд байгаа нь дээрх графикаас харагдаж байна. Дээрх шинжүүдээс бид зөвхөн шүлгийн өгөгдөл болох дараах шинжүүдээр нь машин сургалтын аргаар ангилах боломжийг судалсан үр дүнг туршилтын үр дүн хэсэгт тайлбарлалаа. Онцлог шинж: Уртын дууны өгөгдлөөс дууны нэр(string), бадгийн тоо (numeric), мөрийн тоо (numeric), үгийн тоо (numeric), шүлэг (string), ангиллын нэр (suman, aizam, besreg) зэрэг 5 онцлог (features), 1 ангиллын мэдээлэл (label) ашигласан. Машин сургалтын аргууд нь өгөгдлийн түгэлт, шинж чанараас хамааран харьцангуй үр дүн өгдөг тул weka [18] програмыг ашиглан боломжит аргуудыг туршиж үзсэн. Дараах зурагт weka програмд уншуулах *.arff file-д дууны өгөгдлийг хэрхэн бэлтгэснийг жишээгээр харуулав. Өмнөх хэсэгт өгүүл-

```
@relation 'longsong'
@attribute name string
@attribute badag numeric
@attribute mur numeric
@attribute ug numeric
@attribute Text string
@attribute class {suman,aizam,besreg}
@data
'Urkhan khongor salkhi',12,48,144,'Urkhan khongor salkhind ni \n
Ulias mod gankhana \n Uyakhn zant ter min \n Udakh tusmaa
sanagdana.\n KHukh baglagar modond \n KHukhuu shuvuu dongodono\n
KHuurkhiy uyan setgeliyg \n KHudulgukh yanztay \n KHaliuran
gankhakh zulgend \n KHalj meltrek gold \n KHamag shuvuu
ganganakhad \nKHarin n eldev sanagdana \nYul met orgilson \n
undur Khangayn shield \nungiyn tsetsegs naygaad \nuuriyn nutag
'sanagdana \nKHayrt khunii ur ni \nKHatiraa sayn tsagaanaa
\nKHavtlasan soyj unamaan \nKHajz mordokh yanztay \nKHaz elegt
ulstay \nKHanilj suukhad berkh biy \nTuunee gej sanaval \nTvedej
yuundaa suukh ve \nArvan kheden nastaydaa\nAav eejiynkhee atgand
\n Ider nasaa guytsekhd \nErgej yuundaa khorogdokh ve\nZerleg
burgediyn zulzaga ni \nJiguur suduu guytsekhd \n uujym ugturguyd
nisekhiyg \n Olon bugdeer medne',suman
'Ertniy saykhan',5,20,61,'Ertniy saykhan eruuluur \n Ene saykhan
orond ni \n Ezen Bogd CHingis khaanykhaa \n YAzguuran eets
burdvee khuu \n Tenger met zayaatay \n Tegsh tavan oyuuntay \n
Tenuun uujim zarligtay \n Tengeriyn min ur sad bilee \n Olon odny
dundaas \n Sarny gerel todorkhoy \n Olon khumuuniy dundaas \n
YAndar met chimegtey \n Sumer uulyn oroogoo \n Sunderlej
tsatsaraad \n Sur javkhan delgereed \n Suld n bat boltugay \n
KHamag nomyn rashaanyg \n KHanakhyn chineegeer khurteed \n KHaan
ezen CHingisiyn umnu \n KHamt bugdeeree jargaya',suman
```

Зураг 5: Машин сургалтын таниулах өгөгдлийн формат

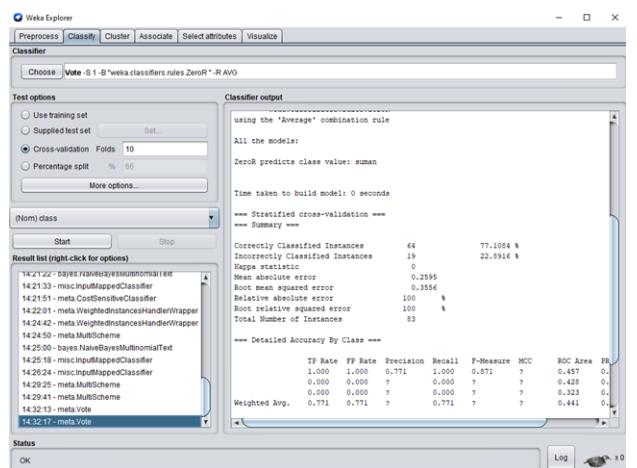
нээр уртын дууны шүлэг нь гүн ухаан, ертөнцийн мөн чанараас гадна хүмүүний аж амьдрал, сэтгэлийн хөдлөлийг тусгасан олон талын утгыг илэрхийлдэг. Тэгвэл цаашдаа эдгээр шүлгийн үгийг машин сургалт ашиглан таниулснаар тухайн дууны утга хийгээд, ерөнхий агуулгыг дүрсээр илэрхийлэх боломж нээгдэх юм. Уртын дууны бүх түрлэгийг нэг дор дуулж, дууны утгыг бүрэн гаргах боломж тааруу, олон түрлэгийг сонсоод суух зав хомс, цагтай уралдан байдаг өнөө цагийн хүмүүст болон уртын дууг шимтэн сонсдог гадна, дотны сонирхогч, судлаач нарт ихээхэн ач холбогдолтой ажил болно гэдэгт итгэлтэй байна. Ингэхийн тулд бид цуг-

луулсан өгөгдөл дээрээ машин сургалтын аргуудыг ашиглан ангилал хийсэн. Зураг 6-д харуулсанчлан



Зураг 6: Weka програм дээр өгөгдлийг уншсан байдал

өгөгдлийн хувьд уртын дууны бүтцийн 3 ангиллаас тэнцүү биш байгаа нь ангиллын үр дүнд нөлөөлж байж болох юм. Ялангуяа бэсрэг төрөл илүү цөөн байна. Эдгээр 3 ангиллаас ихэнх нь суман дуу байдаг. Айзам хэмжээний дуу бусад төрлөөсөө цөөн тоотой байдаг бол бэсрэг дуу ч мөн суман дуунаас хамаагүй цөөн. Машин сургалтын үр дүн ч үүний жишгээр гарч байгаа нь бидний судалгааны үр дүн хүний оюунаар ялгах боломжтой уртын дууны ангиллыг машин сургалтын аргаар ангилах боломж байгааг харуулж байна. Зураг 7-д үзүүлсэн жишгээр



Зураг 7: Weka програм дээрх ангиллын үр дүнгийн жишээ

текстэн өгөгдөлд авч болдог аргуудыг харьцуулж үзсэн ба хамгийн ихдээ 77 хувь буюу 83 дууны 64-ыг зөв ангилсан байна. Энд 10-фолд кросс валидашн зарчмаар шалгасан. Ерөнхийдөө амжилттай хариу үзүүлж байгаа аргууд нь Instance based буюу

сургалтын жишээнүүдийг тестийн жишээтэй харьцуулалт хийж ангилдаг төрлийн аргууд байна. Цаашид энэ чиглэлийн аргуудаар туршихад боломжтой гэж үзлээ. Олон улсад хийгдэж байгаа судалгааны үр дүнг харахад шүлгийн өгөгдөл дээр 50-60хувь [6], аудионы өгөгдөл дээр 80 орчим хувь [8] зөв ангилалтай байсан. Бидний үр дүн 5 онцлог шинжтэй, 3 ангилалд, тэнцвэртэй тооны биш өгөгдөлтэй (unbalanced dataset) үед ийм хэмжээний үр дүн гарч байгаа нь уртын дууны судалгаанд машин сургалтыг хэрэглэх боломж байна гэсэн итгэл төрүүлж байна.

6 Дүгнэлт

Энэхүү судалгааны ажлаар уртын дууны шүлэг, яруу найраг, түүний бүтэц, үгийн утгын зан үйлийн бэлгэдэл зэрэг уртын дууны үгийн онцлог шинжийг судлахын зэрэгцээ уртын дууны уламжлалт судалгааг орчин үеийн технологийн дэвшилтэй хослуулан судлах боломжийг салбар дундын байдалтай авч үзлээ. Уртын дууны шүлгээс өгөгдөл цуглуулан хүний оюун ухаанаар ангилал хийн бэлтгэж, уг өгөгдлөөс машин сургалтын аргуудаар хэрхэн ангилах боломжийг харуулсан нь энэ судалгааны гол үр дүн юм. Аман зохиол судлалын хувьд уламжлалт судалгааны арга түлхүү хэрэглэгдсээр ирсэн. Энэ салбарын судлаачдын судалгааны үр дүнд аман зохиол, өв соёлын арвин сан цугларсан нь бидний энэхүү судалгааны голлох хэрэглэгдэхүүн болж буйд талархалтай байна. Уламжлагдан ирсэн эдгээр сан хөмрөгөө цахим судалгааны эргэлтэд оруулж салбар дундын судалгаагаа өргөжүүлэх боломж бий. Цаашид тус туршилтад сайн ажиллаж буй эдгээр аргуудыг програмчлах, өгөгдлийн шинжилгээг илүү нарийн хийх, шүлгээс дүрс үүсгэх, шүлгийн болон дүрсийн мэдээллийг хамтад нь ашиглах, тодорхой гол түлхүүр үг оноох, ингэснээр олон улсад ч монгол уртын дууны онцлогийг харуулах боломж бүрдэнэ гэж үзэж байна.

Зохиогчийн оролцоо

Н.Баярсайхан уртын дууны шүлгийн судалгааг хийж, ангилал үүсгэхэд шаардагдах онцлог шинжийг тодорхойлж, сургалтын өгөгдлийг бэлтгэж, үр дүнг боловсруулсан. Б.Сувдаа үр дүнгийн үнэн зөв байдлыг хянах, ямар арга алгоритм хэрэглэхийг тодорхойлон чиглүүлсэн. Г.Нандинбилэг уртын дууны судалгааны арга зүй, шинэлэг талыг хянаж зөвлөсөн. Б.Сувдаа өгүүллийг засаж сайжруулах, утга агуулгын алдааг хянаж, сүүлчийн хувилбарыг бичсэн.

Ашиг сонирхлын зөрчилгүйн баталгаа

Зохиогчид ашиг сонирхлын зөрчилгүй гэдгээ баталж байна.

Ашигласан ном

- [1] Алимаа А. Уртын дууны төрөл зүйл, тархалт, хувилбарын онцлог. 2013.
- [2] Галбаяр. Монгол шүлгийн толгой холбох арга, түүний зүй тогтол. ЭШ өгүүллэл. 2018.
- [3] Мэнд-Ооёо. THE INFLUENCE OF ASIAN POETRY ON WORLD LITERATURE. 2019. Available from: <http://www.mend-ooyo.mn/>.
- [4] Сампилдэндэв, Яцковской. Монгол ардын уртын дуу. 1984.
- [5] Дорждавга. Уртын дуу. 1970.
- [6] Tsaptsinos A. Lyrics-based music genre classification using a hierarchical attention network. arXiv preprint arXiv:170704678. 2017.
- [7] Anna Boonyanit ML Andrea Dahl. Music Genre Classification using Song Lyrics. Stanford CS224N Custom Project. 2013.
- [8] Tzanetakis G, Cook P. Musical genre classification of audio signals. IEEE Transactions on speech and audio processing. 2002;10(5):293-302.
- [9] Howard S, Silla Jr CN, Johnson CG. Automatic lyrics-based music genre classification in a multilingual setting. In: Proceedings of the Thirteenth Brazilian Symposium on Computer Music; 2011. .
- [10] Akalp H, Furkan Cigdem E, Yilmaz S, Bolucu N, Can B. Language Representation Models for Music Genre Classification Using Lyrics. In: 2021 International Symposium on Electrical, Electronics and Information Engineering; 2021. p. 408-14.
- [11] Yoon S. Remains and Renewals: The Process of Preserving Urtyn Duu in Contemporary Mongolia. Mongolian Studies. 2013;35:119-31.
- [12] Дулам. Монгол соёл иргэншлийн утга тайлал. 2013.
- [13] Дулам. Бэлгэдэл зүйн судалгаа. 2018.
- [14] Сампилдэндэв. Монгол аман зохиол, зан үйл судлалын бүтээл. 2009.
- [15] Хорлоо. Монгол ардын дууны яруу найраг. 1981.

- [16] Баярсайхан. Төв халхын ургын дууны судалгаа. 2006.
- [17] Unesco. Unesco report. 2008. Available from: <https://ich.unesco.org/en/RL/urtiin-duu-traditional-folk-long-song-00115>.
- [18] Eibe F, Hall MA, Witten IH. The WEKA workbench. Online appendix for data mining: practical machine learning tools and techniques. In: Morgan Kaufmann. Elsevier Amsterdam, The Netherlands; 2016. .