

Компьютерын ухаан

Ангиллагчийн гүйцэтгэлийг үнэлэх хэмжүүр, Өсвөр үеийхний сэтгэл хөдлөлийн өгөгдөл

Т.Ахыт^{1,*}, Ц.Жавзмаа¹, Ч.Баасандорж¹, Л.Ажнай¹, В.Баярмаа²

¹АШУҮИС, Био-Анагаахын Сургууль, Физик Мэдээлэл Зүйн Тэнхим

²Сэтгэцийн Эрүүл Мэндийн Үндэсний Төв

Хүлээн авсан 2023-05-12; Хянагдсан 2023.07.01; Зөвшөөрөгдсөн 2023.07.01

*Холбоо баригч зохиогч: akhit@mnums.edu.mn

Хураангуй

ДЭМБ-ын судалгаагаар дэлхий дахинаа хүүхэд, өсвөр насныхны 20% сэтгэц, зан үйлийн эмгэгтэй ба ихэнхи эмгэг өсвөр насанд эхлэн илэрдэг. Бид энэхүү судалгааны ажлаар Говь-Алтай аймгийн 3764 сурагчдаас онлайнгаар өсвөр насныхны сэтгэл хөдлөл, зан үйлийн илрэлийг үнэлэх SDQ асуумж авч, сурагч, эцэг эх/асран хамгаалагч, багшийн үнэлгээний сангууд үүсгэн өсвөр насныхны сэтгэл хөдлөл, зан үйлийн илрэлийн үнэлгээний өгөгдөлд ангиллын аргуудыг туршин тэдний сэтгэцийн эрүүл мэндийн төлөвийг таамаглахыг зорилоо. Сурагчийн үнэлгээний санд шийдвэрийн модны C50 алгоритм, бэйсийн ангилал, дүрэмд үндэслэсэн ангиллын RIPPER алгоритмаар ангиллагчууд байгуулахад бэйсийн ангиллаар тооцоолсон ангиллагчийн гүйцэтгэлийг үнэлэх хэмжүүр нарийвчлал бусдаас 1-2% их, sensitivity 89% specificity 88% буй нь сурагчдын сэтгэл хөдлөлийн зааг, хэвийн бус илрэлийг жигдхэн ангилсан байна.

Түлхүүр үг: Өгөгдлийн тандалт, ангилал, C50, RIPPER, бэйсийн ангилал, sensitivity, статистик тест

1 Удиртгал

Орчин үед интернэт технологи, харилцаа холбооны хурдацтай ахиц дэвшил нь асар их хэмжээний өгөгдөл бий болгосоор байна. Өгөгдөл олборлолт нь энэхүү хуримтлагдсан их хэмжээний өгөгдлөөс мэдлэг, үнэ цэнтэй мэдээллийг илрүүлэх, нээх үйл явц юм [1–4].

WHO-AIMS-ын үнэлгээгээр “Монгол улсад хүүхдийн сэтгэцийн эрүүл мэндийн тусламж үйлчилгээ байхгүй байна. Цаашид хүүхэд, өсвөр насныханд зориулсан сэтгэцийн эрүүл мэндийн үйлчилгээ явуулах мэргэжлийн чадамжийг хөгжүүлэхэд онцгой анхаарах хэрэгтэй байна.” гэж тэмдэглэжээ.

СЭМҮТ-ийн 2013 оны судалгаагаар өсвөр насны хүүхдүүдийн 40% нь зааг завсрын болон хэвийн бус, эмгэгтэй байна [5].

Монгол улсад өсвөр насанд түгээмэл тохиолддог сэтгэл хөдлөл, зан үйлийн судалгаа хийгдэж байгаа ч их өгөгдөлтэй ажиллах өгөгдлийн шинжилгээний аргаар судлаагүй нь энэ судалгааны ажлын шинэлэг тал болж байна.

Өгөгдөл тандалт нь өвчтөний эрсдэлийг үнэлэх, аливаа өвчин, эмгэгийн төлөвийг урьдчилан таамаглах загварыг бий болгоход туслах маш сайн гүйцэтгэлийг харуулах ач холбогдолтой тул сүүлийн үед анагаах ухааны судалгааны тэргүүлэх салбар болж байна [6]. Тиймээс энэ судалгаа нь эрүүл мэнд болон бусад салбарын судлаачдад их өгөгдөлтэй ажиллахад дэмжлэг үзүүлэх зорилгоор өгөгдөл олборлох

технологийг ашиглах тодорхой бөгөөд ойлгомжтой ойлголттой болоход нь туслах явдал юм.

2 Судлагдсан байдал

Шведийн 7638 ихэр хүүхдийн 474 таамаглагчтай эцэг эхийн тайлан, бүртгэлд машин сургалтын санамсаргүй ой, support vector machine(SVM), мэдрэлийн сүлжээ, XGBoost техникүүдээр өсвөр насны дунд үеийнхний сэтгэцийн эрүүл мэндийн асуудлуудыг урьдчилан таамаглах загвар боловсруулан, AUC хэмжүүрээр логистик регрессстэй харьцуулахад логистик регресссээс бага зэрэг ялгаатай байсан ч итгэлийн интервал нь загвар бүрийн хувьд давхцсан. Санамсаргүй ойн загвар (AUC = 0.739, 95 хувь CI 0.708–0.769), дэмжлэгийн вектор машинууд (AUC = 0.735, 95 хувь CI 0.707–0.764) байна. Загвар нь эмнэлзүйн хэрэглээнд тохиромжгүй боловч сэтгэцийн эрүүл мэндийн ерөнхий үр дүнг урьдчилан таамаглах ирээдүйн загваруудад чухал суурийг тавьсан гэж дүгнэжээ [7].

[8] номонд тодорхойлсон загварууд алдааны дундаж түвшиндээ бодит ялгаатайг тодорхойлох статистикийн ач холбогдлын тестийг энэ судалгааны ажилд байгуулсан загваруудад туршин, хоорондоо хэр ялгаатайг тооцоолоход бэйсийн ангиллагч 2 хувь бага зэрэг илүү үр дүн үзүүлсэн.

Олон улсын судалгааны ажлуудад [9–11] ангиллагчийн гүйцэтгэлд нарийвчлал, мэдрэмж, онцлог

байдал, F1- хэмжүүрүүдийг загваруудын үнэлгээнд ашиглаж байна. Харин загварууд ялгаатайг тодорхойлох статистикийн ач холбогдлын тестийг эрүүл мэндийн өгөгдлөөс үүссэн загварт туршсан судалгааны ажил маш ховор байгаа нь энэхүү судалгааны ажлын шинэлэг тал болж байна.

3 Арга зүй

3.1 Хүчтэй ба сул талын тухай асуумж (SDQ)

Олон улсад түгээмэл хэрэглэгддэг өсвөр насныхны сэтгэл хөдлөл, зан үйлийн илрэлийг үнэлж, эмгэгийг илрүүлэх 25 асуулт, 3 шалгууртай SDQ асуумжийг хүүхэд, эцэг эх/асран хамгаалагч, ангийн багшаас тус тусад нь авч ашигласан. Асуумж нь өсвөр үеийнхний сэтгэл зүйн үнэлгээний 25 шалгуураас гадна ерөнхий, хүн ам зүйн үзүүлэлт, сэтгэл зүйн бэрхшээлийн, нөлөөлөх хүчин зүйл зэрэг бүлгээс бүрдэх бөгөөд сэтгэл хөдлөл, төрх үйл, анхаарал дутмагшил/хэт хөдөлгөөнтөх, үе тэнгийг харилцаа, нийтэч занг тус тус илрүүлнэ. Хүүхдийн хариултаар дотоод шинж чанар сэтгэл хөдлөлийн “Нийт оноо” 0-5 үед хэвийн, 6 байхад зааг, 7-10 үед хэвийн бус илрэлтэй гэж үзнэ [5, 12].

Энэ ажлаар SDQ асуумжийг онлайнаар авч, өгөгдөл цуглуулан өгөгдлийн тандалтын аргаар өсвөр үеийнхний сэтгэцийн ерөнхий төлөв байдлыг илрүүлэх зорилготой юм.

3.2 Ангиллын аргууд

Ангиллын арга нь багцаас чухал ач холбогдолтой ангиллыг илэрхийлэх загвар олборлох өгөгдлийн шинжилгээний нэг хэлбэр юм. Шийдвэрийн мод, бейсийн ангилал, дүрэмд үндэслэн (rule-based) зэрэг үндсэн аргууд байдаг [8].

Шийдвэрийн мод нь класс хаягтай сургалтын багцаас шийдвэрийн мод сургах ажиллагаа бөгөөд зангилаа бүр шинж чанар, салаа туршилтын үр дүн, навч класс хаягийг (class label) тодорхойлох шаталсан хэлбэртэй мод бүтэц болно [13].

Шийдвэрийн модны CHAID, C50, CART зэрэг олон алгоритм байдаг. C50 нь ID3 алгоритмаас хурд, санах ойн ашиглалт, жижиг мод, олон ангиллыг жигнэх, ашиггүй атрибутуудыг автоматаар ялгах зэрэг олон давуу талуудтай [13].

Бейсийн ангилал (naive Bayes classifiers) нь шинж чанаруудын хоорондын бие даасан байдлыг Бейсийн теоремд үндэслэн таамагладаг статистик магадлалуудын нэг бөгөөд шийдвэрийн мод, мэдрэлийн сүлжээ ангиллагчидтай гүйцэтгэлийн хувьд харьцуулах боломжтой, томоохон мэдээллийн сангуудад өндөр нарийвчлал, хурдыг үзүүлдэг [8].

Дүрэмд үндэслэсэн (rule-based) ангилал нь мэдлэгийг дүрслэх шилдэг аргуудын нэг бөгөөд загвар нь урьдчилсан нөхцөл (rule antecedent) IF condition,

AND холбоосоор бүрдсэн атрибутуудын туршилтаас, дүрмийн үр дүн (rule consequent) THEN класс таамаглалуудын иж бүрдлээр илэрхийлэгддэг. Хэрвээ урьдчилсан нөхцөл нь өгөгдсөн X багцын хувьд үнэн байвал дүрэм багцыг хамарч, үр дүн нь хүчин төгөлдөр байна. D өгөгдлийн иж бүрдэл дахь класс анги бүхий X багцыг ашиглан R дүрмээр хамрагдсан бичлэгийн тоо n_{covers} , R дүрмээр зөв ангилагдсан бичлэгийн тоогоор $n_{correct}$ хамрах хүрээ (coverage), нарийвчлалыг (accuracy) тооцож үнэлдэг. Хэмжээгээр дараалуулах (size ordering), дүрмээр дараалуулах (rule ordering) зэрэг олон арга байдаг. Дүрмээр дараалуулах нь AQ, CN2 ба RIPPER зэрэг өргөн хэрэглэгддэг алгоритмуудтай [8].

R программын $train(x, y, \dots)$ функц нь шийдвэрийн мод, бейсийн ангилал, дүрэмд суурилсан (rule-based) ангиллаар загварууд тооцоолдог [14, 15].

3.3 Загварын үнэлгээ

Ангиллын аргуудаар загварууд байгуулан, тэдгээрийг хооронд нь гүйцэтгэлийг үнэлэх хэмжүүрээр харьцуулан, оновчтой загвараар мэдлэгийг дүрсэлдэг. Иймд ангиллагчийн урьдчилан таамаглах чадварыг үнэлэх нарийвчлал (accuracy), мэдрэмж (sensitivity), онцлог байдал (specificity), статистик ач холбогдлын шалгуур (statistical test of significance STS) гэх мэт үнэлгээний хэмжүүрүүдийг судаллаа [8].

Ангиллын аргуудыг ашиглан сургалтын өгөгдлөөр загварууд үүсгэн, туршилтын өгөгдлөөр загваруудын нарийвчлал, мэдрэмж, онцлог байдал, статистик ач холбогдлын шалгуур зэрэг хэмжүүрийг тооцон, гүйцэтгэлийг үнэлж харьцуулан, оновчтойг нь сонгоно.

Нарийвчлал нь ангиллагчаар зөв ангилагдсан туршилтын багцын хувь, мэдрэмж нь зөвөөр хаягласан эерэг багцын хувь, онцлог байдал нь зөвөөр ангилсан сөрөг багцын хувь болно. Өгөгдлийн багцын тархалт нь сөрөг ангилал дийлэнх олонх, цөөнх нь эерэг ангилалтай буюу жигд бус байж болно. Энэ үед эерэг болон сөрөг багцыг зөв ангилж ялгах sensitivity, specificity хэмжүүрийг ашигладаг [8].

Хөндлөн баталгаажуулалтаар өгөгдлийг арав хуваан ангиллын(10-fold cross validation) аргуудаар M1, M2, M3 загвар байгуулан загвар бүрийн дундаж алдааны түвшнийг тооцон хамгийн сайныг тодорхойлохын тулд алдаа багатай загварыг сонгож болно. M1, M2, M3 загваруудын алдааны дундаж түвшин ялгаатай байх ч статистик ач холбогдолгүй, ялгаа нь тохиолдлоос үүдэлтэй байж болно [8].

Загварууд алдааны дундаж түвшиндээ мэдэгдэхүйц ялгаатайг тодорхойлохын тулд STS буюу стьюдентийн t шалгуурыг (t-test, or Student's t-test) тооцоолж, $sig=5\%$ ач холбогдлын түвшнийг сонгон, тархалтын хүснэгт дэх k - 1 эрх чөлөөний зэргийн $sig/2$ -д харгалзах z утгатай харьцуулснаар бодит ял-

гааг тодорхойлон "Нэг загвар нь илүү дээр" мэдэгдлийг хийх боломжтой юм [8].

Хөндлөн баталгаажуулалтаар сургалтын өгөгдлийг 10 хуваан ангиллын аргуудаар M_1, M_2, M_3 загварууд үүсгэн, загвар бүрээр олж авсан 10 алдааны түвшнүүдийн дунджийг тооцож тэдгээрийн хоорондох алдааны дундаж түвшний зөрүү тэг гэсэн таамаглалыг үгүйсгэн, статистикийн хувьд мэдэгдэхүйц ялгаатайг тодорхойлно. Үүний тулд ач холбогдлын STS буюу стьюдентийн t (t -test, or Student's t -test) шалгуураар тэг таамаглалыг шалгана. Өгөгдсөн загварын хувьд хөндлөн баталгаажуулалтад тооцсон алдааны хувь хэмжээнүүдийг бие даасан түүвэр гэж үзэн $k-1$ зэрэгтэй бөгөөд $k=10$ байна. Хэрэв тэг таамаглалыг үгүйсгэж чадвал загваруудын ялгаа нь статистикийн хувьд ач холбогдолтой гэж дүгнэж алдааны түвшин багатай загварыг сонгоно [8].

$$t = \frac{\overline{\text{err}}(M_1) - \overline{\text{err}}(M_2) - \overline{\text{err}}(M_3)}{\sqrt{\text{var}(M_1 - M_2 - M_3)/k}} \quad (1)$$

$$\text{var}(M_1 - M_2 - M_3) = \frac{1}{k} \sqrt{\text{err}(M_1)_i - \text{err}(M_2)_i - \text{err}(M_3)_i} \quad (2)$$

4 Туршилт

Говь-Алтай аймгийн 3764 сурагч, 3176 эцэг эх асран хамгаалагч, анги даасан багш нараас онлайнгаар SDQ асуумж авч, 10122 өгөгдөл цуглуулан, ерөнхий, хүн ам зүйн, сэтгэл хөдлөл, зан үйл, анхаарал дутмагшил/хэт хөдөлгөөнтөх, үе тэнгийнхний харилцаа, нийтэч зангийн илрэлийг үнэлэх, бэрхшээл, нөлөөлөх хүчин зүйл, сэтгэл хөдлөл, зан үйлийн нийт үзүүлэлт гэх мэт атрибуттай сурагч, багш, асран хамгаалагчийн үнэлгээний сангууд үүсгэлээ. Үүнээс сурагчдын үнэлгээний санд дотоод шинж чанар болох сэтгэл хөдлөл, үе тэнгийнхний харилцааны үзүүлэлтүүдээс сэтгэл хөдлөлийн хэвийн, зааг, хэвийн бус төлөв хэрхэн хамаарахыг ангиллын аргуудаар судлав. Хүснэгт 1-д сурагчийг үнэлэх сэтгэл хөдлөл, үе тэнгийнхний харилцааны асуулт болон оноог, Зураг 1-т сурагчдын үнэлгээний сангийн харагдах байдлыг харууллаа.

Сурагчдийн өөрийгөө үнэлэх сангийн зааг 1355 болон хэвийн бус төлөвийн 1736, нийт 3091 бичлэгийг 10 хуваасан хөндлөн баталгаажуулалтын аргаар загвар, нарийвчлалыг тооцон, Зураг 2-т үзүүлэв.

Загваруудын хуваалт бүрийн алдааны түвшнийг тооцон, $\text{err}(M_1)_i$, $\text{err}(M_2)_i$, $\text{err}(M_3)_i$ тооцон $\overline{\text{err}}(M_1)$, $\overline{\text{err}}(M_2)$, $\overline{\text{err}}(M_3)$ дунджилсан. Хуваалт бүрд тооцсон алдааны хувь хэмжээг бие даасан түүвэр гэж үзэн, дисперсийг $\text{var}(M_1 - M_2 - M_3)$ олсон. k дээжийн хувьд $k-1$ эрх чөлөөний зэрэгтэй t статистикийг тооцдог. Бидний судалгааны ажлын хувьд $k=10$ бөгөөд ач холбогдлын тестийг шалгах

S3	S8	S13	S16	S24	S6	S11	S14	S19	S23	InterCI
0	2	0	1	0	1	2	1	0	0	2
2	2	0	1	0	1	0	0	0	0	2
1	1	1	0	1	0	1	1	0	0	2
1	0	1	1	0	2	1	1	0	1	2
0	2	0	1	0	0	2	1	0	1	2
0	1	0	1	0	1	2	1	0	0	2
0	1	0	1	1	0	2	0	0	0	1
0	1	1	1	1	0	2	0	0	0	2
1	1	1	1	1	2	2	1	0	0	3
1	2	1	1	1	0	2	1	0	0	3
0	1	1	1	1	0	2	2	1	2	3

Зураг 1: Сурагчдын үнэлгээний сангийн харагдах байдал

Хүснэгт 1: Сурагчдын дотоод шинж чанарын үзүүлэлт

№	Атрибут/ асуултууд	Үнэн биш	Зарим -даа	Яг үнэн
Сэтгэл хөдлөл				
1	Ихэнхдээ толгой өвддөг (S3)	0	1	2
2	Олон зүйлд санаа зовнидог, түгшдэг (S8)	0	1	2
3	Ихэнхдээ аз жаргалгүй, сэтгэлээр унадаг (S13)	0	1	2
4	Шинэ орчинд байхаараа сандарч түгшдэг (S16)	0	1	2
5	Олон зүйлээс айдаг (S24)	0	1	2
Үе тэнгийнхний харилцаа				
1	Ихэнхдээ ганцаараа тоглодог (S6)	0	1	2
2	Нэг л дотны сайн найзтай (S11)	2	1	0
3	Нэг л дотны сайн найзтай (S14)	2	1	0
4	Бусад хүүхдүүд дарамталдаг (S19)	0	1	2
5	Үеийн хүүхдүүдтэй байхаасаа илүү том хүмүүстэй харьцах нь амар байдаг (S23)	0	1	2

боломжийг олгодог. M_1, M_2, M_3 хоорондоо мэдэгдэхүйц ялгаатайг тодорхойлох ач холбогдлын шалгуур t утга нь тархалтын хүснэгтийн хувьд 2 ач холбогдлын түвшинд харгалзах тархалтын $z=-2.26$ утгатай харьцуулахад $t > z$ эсвэл $t < -z$ ханган, тархалтын сүүл хэсэгт, татгалзах мужид оршиж байна. Судалгааны ажлын үр дүнд тооцоолсон STS буюу $t=-9.31$ нь тархалтын хүснэгт дэх $k - 1$ эрх чөлөөний зэргийн $\text{sig}/2$ -д харгалзах $z=-2.26$ утгаас бага, $t < -z$ нөхцөлийг хангаж тэг таамаглалыг няцаав. Энэ нь загварууд хоорондоо статистикийн хувьд мэдэгдэхүйц ялгаатай тул алдааны түвшин бага бейсийн ангиллын аргаар тооцоолсон загварыг сонгох боломжтойг Зураг 3-аар харуулж байна.

Судалгаанд хамрагдсан сурагчдын үнэлгээгээр

		Таамагласан			Accuracy
		2	3		
C50 – M ₁	2	36.7	6.6	0.8620	
	3	7.2	49.5		
Naïve Bayes -M ₂	2	39	6.5	0.8869	
	3	4.8	49.6		
RIPPER-M ₃	2	39	6.5	0.8750	
	3	4.8	49.6		

Зураг 2: Загваруудын confusion matrix, нарийвчлал (accuracy)

№	Загваруудын хуваалт тус бүрийн алдаа		
	C50 $err(M_1)_i$	Bayes $err(M_2)_i$	RIPPER $err(M_3)_i$
1	0.16	0.10	0.12
2	0.14	0.12	0.14
3	0.11	0.13	0.12
4	0.10	0.10	0.13
5	0.14	0.11	0.12
6	0.12	0.14	0.13
7	0.15	0.10	0.16
8	0.14	0.12	0.13
9	0.15	0.12	0.08
10	0.16	0.10	0.12
\overline{err}	0.14	0.11	0.13
$\overline{err}(M_1) - \overline{err}(M_2) - \overline{err}(M_3)$			-0.10
$\sqrt{var(M_1 - M_2 - M_3)/k}$			0.01074547619
Z итгэх хязгаар			-2.26
T шалгуур			-9.31

Зураг 3: Бэйсийн ангиллын аргаар тооцоолсон загварын үр дүн

Хүснэгт 2: Сурагчийн дотоод шинэ чанарыг илэрхийлэх загвар

State	0	1	2	0	1	2
Сэтгэл хөдлөл				Үе тэнгийнхний харилцаа		
S3				s6		
2	0.69	0.28	0.03	0.64	0.30	0.04
3	0.33	0.52	0.15	0.31	0.505	0.18
s8				S11		
2	0.18	0.62	0.19	0.06	0.16	0.77
3	0.03	0.46	0.5	0.05	0.22	0.71
s13				s14		
2	0.69	0.25	0.015	0.11	0.55	0.33
3	0.23	0.60	0.16	0.08	0.61	0.30
s16				s19		
2	0.32	0.55	0.12	0.80	0.17	0.01
3	0.09	0.51	0.38	0.46	0.43	0.10
s24				s23		
2	0.53	0.40	0.05	0.51	0.44	0.03
3	0.13	0.54	0.31	0.28	0.57	0.14

сэтгэл хөдлөлийн зааг илрэлийн магадлал заримдаа олон зүйлд санаа зовниоход 0.62, аз жаргалгүй, сэтгэлээр унахдаа 0.6, шинэ орчинд сандарч, түгшихдээ 0.55, олон зүйлээс айхад 0.55 байна. Харин үе тэнгийнхний харилцааны зааг илрэлийн магадлал нь ганцаараа тоглохдоо 0.5, хэвийн бус илрэлийн магадлал нэг дотно найзтай үедээ 0.7, заримдаа бусад хүүхдүүдэд таалагдахгүй байхад 0.62, үе, тэнгийнхний хүүхдүүдтэй байхдаа 0.57 байна (Хүснэгт 2).

Сурагчдын үнэлгээний санг 10 хуваасан хөндлөн баталгаажуулалтын аргаар нарийвчлал нь 70:30 харьцаагаар сургалт, туршилтын багцад хоёр хувааснаас 10 % их, 0.86-0.89 хооронд байгаа нь хөндлөн баталгаажуулалт өгөгдөл их байх тусам загварын нарийвчлал өндөр хувьтай байна.

5 Дүгнэлт

- Өгөгдөл их байх тусам хөндлөн баталгаажуулалт ашигласан ангиллын загваруудын нарийвчлал, мэдрэмж, онцлог байдал зэрэг нь өндөр хувьтай байх төлөвтэй байна.
- Статистик ач холбогдлын шалгуураар загварууд нь хоорондоо статистик ялгаатай учраас алдаа бага бэйсийн ангилал нь өгөгдөл их үед сайн ажиллах төлөвтэйг илтгэж байна.
- Судалгааны үр дүнгээс харахад анагаах, эрүүл мэндийн өгөгдөлд ангиллын аргуудаар загвар байгуулан, гүйцэтгэлийн хэмжүүрүүдээр үнэлэхэд нэлээд өндөр хувьтай гарч байгаа нь цаашид судлаачид ашиглах боломжтойг харуулаа.
- Судалгаанд хамрагдсан сурагчдын үнэлгээгээр тэдний сэтгэл хөдлөлийн зааг илрэл нь заримдаа олон зүйлд санаа зовниох, аз жаргалгүй, сэтгэлээр унах, шинэ орчинд сандарч, түгшихэд, олон зүйлээс айхад, харин үе тэнгийнхний харилцааны зааг илрэл нь ганцаараа тоглохдоо, хэвийн бус илрэл нь нэг дотно найзтай байх, заримдаа бусад хүүхдүүдэд таалагдахгүй байхад, үе, тэнгийн хүүхдүүдтэй байхдаа үүсэх төлөвтэй байна.

Зохиогчийн оролцоо

В.Баярмаа нь уг судалгааны ажлын санааг гаргасан бол өгөгдөл цуглуулах, ерөнхий загварын архитектур, туршилт болон дүн шинжилгээг бүх зохиогч нар ижил оролцоотой гүйцэтгэсэн болно. Ц.Жавзмаа нь өгүүллийн эхний хувилбарыг бичсэн бол Т.Ахыт өгүүллийн утга, агуулгын алдааг засаж сүүлийн хувилбарыг бичсэн.

Ашиг сонирхлын зөрчилгүйн баталгаа

Бүх зохиогчид ашиг сонирхлын зөрчилгүй болохыг баталж байна.

Ашигласан ном

- [1] Soumen Chakrabarti ME. Data Mining Curriculum: A Proposal (Version 1.0). Intensive Working Group of ACM SIGKDD Curriculum Committee. 2006:1-10.
- [2] Trevor Hastie RT. The elements of statistical learning: data mining, inference, and prediction. Springer. 2009;2.
- [3] Han J, Kamber M. Data Mining Concepts and Techniques. imprint of Elsevier. 2006:559 569.
- [4] Arockiam L, Charles S. Deriving Association between Urban and Rural Students Programming Skills. International Journal on Computer Science and Engineering. 2010:687 690.
- [5] Vaanchindorj B. Өсвөр насныхны сэтгэл хөдлөл, зан үйлийн судалгаа. Ulaanbaatar city. 2019.
- [6] Wu WT, Li YJ. Data mining in clinical big data: the frequently used databases, steps, and methodological models. Military Medical Research. 2021:1-12.
- [7] Tate AE, McCabe RC, Larsson H, Lundström S, Lichtenstein P, Kuja-Halkola R. Predicting mental health problems in adolescence using machine learning techniques. PloS one. 2020;15(4):e0230389.
- [8] Han J, Kamber M, Pei J. Data mining concepts and techniques third edition. University of Illinois at Urbana-Champaign Micheline Kamber Jian Pei Simon Fraser University. 2012.
- [9] Raschka S. An overview of general performance metrics of binary classifier systems. arXiv preprint arXiv:14105330. 2014.
- [10] Hossin M, Sulaiman MN. A review on evaluation metrics for data classification evaluations. International journal of data mining & knowledge management process. 2015;5(2):1.
- [11] Totsch N HD. Classifier uncertainty: evidence, potential impact, and probabilistic treatment. PeerJ Comput Sci. 2021:398-402.
- [12] Goodman R. The Strengths and Difficulties Questionnaire (SDQ), in Innovations in clinical practice: Focus on children and adolescents. Professional Resource Press/Professional Resource Exchange, Inc p. 2003:109-11.
- [13] Palmov SV, Miftakhova AA. Comparison of classification algorithms C4. 5 and C5. 0. Infokommunikacionnye tehnologii. 2015:467-71.
- [14] Max Kuhn SW Jed Wing. Package ‘caret’. The R Journal. 2020:223.
- [15] Max Kuhn SW Jed Wing. A Short Introduction to the caret Package’. R Found Stat Comput. 2015:1-10.