

# Орон сууцны үнийг таамаглахад сургуулийн байршлын нөлөөллийг машин сургалтын арга ашиглан судлах нь

Жамбалдоржийн Отгончимэг<sup>1</sup>, Болдын Наранчимэг<sup>1,\*</sup>

<sup>1</sup>МУИС, МТЭС, Мэдээлэл, компьютерын ухааны тэнхим

Хүлээн авсан 2024.05.08; Хянагдсан 2024.11.19; Зөвшөөрөгдсөн 2024.12.13

\*Холбоо баригч зохиогч: naranchimeg@num.edu.mn

## Хураангуй

Өнөөгийн өрсөлдөөнт зах зээлд орон сууцны үнэ цэнийг тодорхойлохдоо дан ганц зардалд тулгуурлан үнэлэхээс гадна барилгын бүтээц, өрөөний зохион байгуулалт, байрлал, хэрэглэгчийн амьдралын хэв маяг, санхүүгийн үнэ цэн гэх мэт бусад хүчин зүйлсийг харгалзан үздэг болсон. Эдгээр хүчин зүйлсийн нэг нь сургууль юм. Олон улсад сургуулийн чанар болон байршил нь орон сууцны үнэд нөлөөлж буй олон судалгаанууд байдаг. Харин манай улсын хувьд орон сууцны үнийг таамаглах судалгааны ажил тэр дундаа сургуулийн байршлын нөлөөг судалсан ажил байдаггүй тул бид энэхүү ажлаар сургуулийн байрлал Улаанбаатар хотын орон сууцны үнэд хэрхэн нөлөөлж буйг машин сургалтын арга ашиглан судлахыг зорилгоо. Тодруулбал бид үл хөдлөхийн мэргэжлийн байгууллагаас хүлээн авсан 61106 өрхийн, 418 төслийн орон сууцны өгөгдөл дээр ойролцоох 242 улсын сургууль, 65 хувийн сургуулийн байршлын координатыг цуглуулж, хамгийн ойр сургууль хүртэлх зайг тооцоолсон. Ингээд үл хөдлөх хөрөнгийн мэргэжлийн байгууллагаас авсан 25 онцлог шинж чанар бүхий өгөгдөл дээр нэмэлтээр сургуулийн зай, замын хөдөлгөөний бүсчлэл, инфляц, барилгын зардлын индекс зэрэг өгөгдөл цуглуулж түгээмэл хэрэглэгддэг машин сургалтын аргуудыг ашиглан орон сууцны үнийг таамагласан. Туршилтын үр дүнгээс харахад онцлог шинж чанарууд нэмэгдэхэд алдааны үзүүлэлт багассан бөгөөд ялангуяа сургуулийн зайн мэдээлэл орон сууцны үнийг тооцоолоход чухал ач холбогдолтой болох нь батлагдсан. Машин сургалтын аргуудын хувьд XGBoost болон Random Forest аргууд нь хамгийн сайн гүйцэтгэлийг үзүүлсэн.

**Түлхүүр үг:** орон сууцны үнэ; машин сургалт; XGBoost; random forest; сургуулийн байршил

## 1 Удиртгал

Монгол улсын нутаг дэвсгэрт оршин суугаа хүн амын 46.9 хувь нь нийслэл Улаанбаатар хотод оршин суудаг ба нийслэлийн хүн ам сүүлийн 5 жилд тогтмол өсч байна<sup>1</sup>. Үүнтэй зэрэгцээд нийслэлийн орон сууцны эрэлт, нийлүүлэлт нэмэгдэж орон сууц худалдан авах, худалдаалах хэрэгцээ их байна. Орон сууцны үнэд барилгын насжилт, байршил, худалдаа үйлчилгээ, барилгын бүтээц, ногоон байгууламж, барилгын зэрэглэл, бүсчлэл, зогсоол, ойрхон сургууль, цэцэрлэг гэх мэт хүчин зүйлс голлон нөлөөлдөг<sup>2</sup>. Тэр дундаа дэд бүтцийн төлөвлөлт, үл хөдлөх хөрөнгийн зохион байгуулалтгүйгээс болж хотын авто замын хөдөлгөөн, түгжрэл маш ихтэй Улаанбаатар хотод амьдрах орчинтой ойр сургууль байх нь цаг, мөнгө хэмнэх давуу талтай. Үр дүнд нь Улаанбаатар хотод орон сууцны олон хорооллууд баригдсаар<sup>3</sup> байгаа боловч сургуулийн барилга харьцангуй бага баригдаж буй нь сургуулийн хүрэлцээ, хүртээмжийг хумьж сургуулийн эрэлт хэрэгцээг өгсөж байна.

Олон улсад орон сууцны үнэд сургуулийн чанар болон байршил хэрхэн нөлөөлж буйг судалсан олон ажлууд байдаг. Тухайлбал, сургуулийн чанар Парис(1) болон Бээжин(2) хотын орон сууцны үнэд хэрхэн нөлөөлж буйг судалсан байдаг бол Сах болон түүний баг (3) Сан диэго хотын 2010-2011 оны 20000 орон сууцны үнэд сургууль ойр байх нь хэрхэн нөлөөтэй болохыг хедоник тооцоолол ашиглан судалсан байна. Цаашлаад Австралийн Сидней хотын судалгаанд "Орон сууцны үнэд улсын сургуулиудын үзүүлэх нөлөө"(4), Хятад улсын Шанхай хотын "Сургуулийн байршил байрны үнэд үзүүлэх нөлөө"(5) зэрэг орон сууцны үнийг таамаглах ажлуудыг олон улсад хийсэн байна.

Манай улсын хувьд орон сууцны үнийг урьдчилан таамаглах судалгаа харьцангуй цөөн байна. Үүний шалтгаан нь орон сууцыг ихэвчлэн [mongolbank.mn](https://www.mongolbank.mn/), фэйсбүүк групп гэх мэт цахим худалдааны сайтыг ашиглан худалдаалах ба эдгээр нь ямар нэгэн хяналтгүй, богино хугацаанд хандах боломжтой зэрэг сул талтай. Түүнчлэн хүний хэрэгцээ, тухайн бүс нутагт байгаа сургууль, цэцэрлэг зэрэг мэдээллийг худалдан авагч өөрөө судлах эсвэл үл хөдлөх хөрөнгийн экспертийн тусламжтайгаар тодорхойлж байна.

Иймээс бид өмнөх судалгааны ажлаар (6) үл хөд-

<sup>1</sup><https://www.1212.mn/>

<sup>2</sup><https://www.mongolbank.mn/>

<sup>3</sup><https://www.barilga.mn/>

лөх хөрөнгийн мэргэжлийн байгууллагаас 2019 онд Улаанбаатар хотод ашиглалтанд орсон шинэ орон сууцны өгөгдөл хүлээн авч, түүнийг цэвэрлэн боловсруулж машин сургалтын аргуудыг харьцуулан судалж орон сууцны үнийг таамаглахад XGBoost, Random forest машин сургалтын аргууд нь 87% хүртэлх үр дүнг үзүүлсэн.

Иймд бид цаашлаад энэхүү судалгааны ажлаар Улаанбаатар хотын орон сууцны үнэд сургуулийн байршил хэрхэн нөлөөлж буйг машин сургалтын арга ашиглан судлахыг зорилоо. Түүнчлэн барилгын зардлын индекс, замын хөдөлгөөний бүсчлэл, инфляци зэрэг онцлог шинжүүд нь орон сууцны үнэд хэрхэн нөлөөлж буйг судална. Зураг 1-т бидний аргын ерөнхий бүдүүвчийг үзүүлэв. Өмнөх судалгаануудаас харахад сургуулийн байршлын нөлөөллийг машин сургалтын арга ашиглан судалсан судалгаа цөөн, Монгол улсын хувьд анх удаа хийгдэж байгаа гэдгээрээ энэхүү ажил нь шинэлэг юм.

Энэ өгүүллийн хоёрдугаар бүлэгт орон сууцны үнэд сургуулийн байршил нөлөөлж буйг ижил төстэй ажлын талаар тайлбарлана. Гуравдугаар бүлэгт судалгааны арга зүй, өгөгдлийн багцыг хэрхэн бүрдүүлснээ тайлбарлах болно. Харин дөрөвдүгээр бүлэгт орон сууцны үнийг таамаглахад сургуулийн байршил болон нэмэлт онцлог шинжүүд хэрхэн нөлөөлж буйг судлах туршилт, түүний үр дүнг тайлагнаж, эцэст нь судалгааны ажлын дүгнэлтийг танилцуулна.

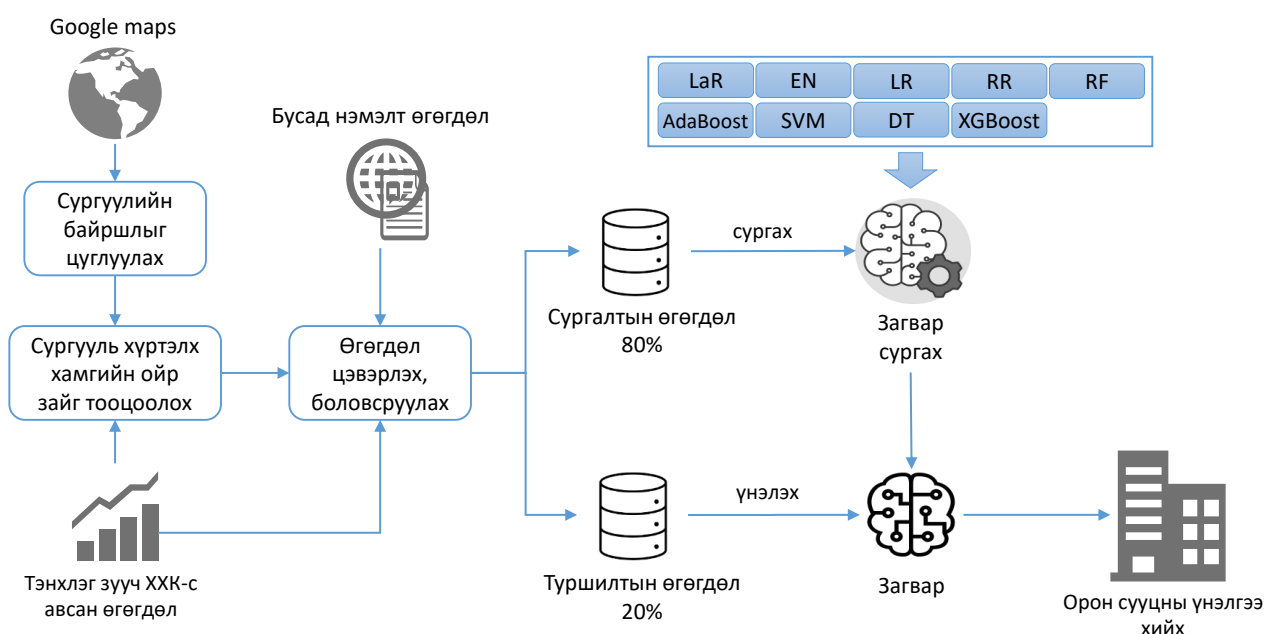
## 2 Судлагдсан байдал

Орон сууцны үнийг урьдчилан таамаглах олон тооны судалгаа сүүлийн жилүүдэд хийгдсэн байдаг ба

энэ хэсэгт бид орон сууцны үнийг машин сургалтын арга ашиглан таамаглах, сургуулийн чанар болон байршил нь орон сууцны үнэд хэрхэн нөлөөлж буйг судалсан судалгаанууд, тэдгээрийн үр дүнг онцлох болно.

Орон сууцны үнийг тодорхойлоход олон хүчин зүйлс нөлөөлж байдаг нь энэ салбарт олон судлаачдын машин сургалт болон гүнзгий сургалтын загварууд боловсруулах сонирхлыг татаж байна(7; 8). Тухайлбал, Хонг Конгийн үл хөдлөх хөрөнгийн салбарт хийсэн судалгаанд Extra trees (ET), k-Nearest Neighbors (KNN), Random forest (RF) машин сургалтын аргыг ашигласан ба үр дүнг уламжлалт арга буюу хедоник үнийн загвартай харьцуулсан судалсан байна (9). Туркийн үл хөдлөх хөрөнгийн салбарт хийсэн судалгаанд 70360 орон сууцны гүйлгээний бүртгэлийг ашиглан Linear regression (LR), Lasso regression (LaR), Ridge regression (RR) перцепс, XGBoost, Artificial neural network (ANN) аргуудыг харьцуулан үл хөдлөх хөрөнгийн үнийг таамаглах судалгаа хийсэн байна(10).

Цаашлаад,(4; 11; 12) орон сууцны үнэд сургуулийн чанар болон байршил хэрхэн нөлөөлж буйг судалжээ. Тухайлбал судлаачид(11) БНХАУ-ын Шанхай хотын орон сууцны үнэд бага сургуулиуд хэрхэн нөлөөлж байгааг судалсан. Тодруулбал, уг судалгаанд өндөр чанартай бага сургууль болон энгийн бага сургуультай ойр орон сууцны үнийн ялгааг судалжээ. Үр дүнгээс харахад өндөр чанартай бага сургуулиуд бүс нутгийн орон сууцны үнэд нөлөө үзүүлж байгаа нь тогтоогджээ. Түүнчлэн чанартай сургуульд ойролцоо байрлах жижиг талбайтай байшингийн үнэ нь хол байрлах их талбайтай байшингийн үнээс өндөр байсан байна(13). Австралийн Сидней хотын 2019 оны 2,3-р улиралд худал-



Зураг 1: Үл хөдлөх хөрөнгийн үнийг таамаглах аргын ерөнхий бүдүүвч

даалагдсан 13534 орон сууцны өгөгдлийг ашиглан орон сууцны үнэд улсын сургуулиудын нөлөөллийг судалсан (4). Судлаачид орон сууцны үндсэн (газрын хэмжээ, өрөөний тоо, ариун цэврийн өрөөний тоо гэх мэт) мэдээллээс гадна орон зайн мэдээлэл (хотын төвөөс хэдэн км, хамгийн ойрын төмөр замын өртөө хүртэлх зай гэх мэт), хөршийн мэдээлэл (мэргэжлийн ажилчдын хувь, гадаадад төрсөн оршин суугчдын хувь, гэр бүлийн орлого гэх мэт) болон харьяа улсын сургууль хүртэлх зай гэх мэт өгөгдлийг ашиглан гурван хедоник регрессийн загварыг харьцуулан судалсан байна. Үр дүнгээс харахад Сиднейн хойд болон баруун бүс нутагт байрлах орон сууцад улсын сургуулиуд хамгийн их нөлөөлдөг гэсэн үр дүн гарсан байна.

Монгол улсын хүрээнд хийгдсэн судалгаанууд нь ерөнхийдөө маш бага ажиглалт, онцлог шинж чанартай байдаг ч боловсруулсан загваруудын төрөл нь нэлээд хязгаарлагдмал. Ихэнх судалгааг орон сууцны тодорхой дэд зах зээлийн түүвэр болгон нэгэн төрлийн мэдээллийн багц дээр хийсэн бөгөөд боловсруулсан загварууд нь нийтэд хүртээлтэй биш байдаг<sup>4</sup>. Бидний судалснаар манай улсад орон сууцны үнийг урьдчилж таамаглах судалгаа маш цөөн ба орон сууцтай ойролцоох улсын болон хувийн сургуулийн байршлыг судалсан судалгааны ажил хийгдээгүй байна.

### 3 Арга зүй

Бидний аргын ерөнхий бүдүүвчийг Зураг 1-т үзүүлэв. Бид орон сууцны үнийг таамаглахдаа үл хөдлөх хөрөнгийн мэргэжлийн байгууллага болох "Тэнхлэг зууч"ХХК-аас Улаанбаатар хотын 2019 онд ашиглалтад орсон 418 төслийн 61108 өрхийн өгөгдөл авсан. Дээр нь орон сууцтай ойролцоох 242 улсын сургууль, 65 хувийн сургуулийн байршлын координатыг цуглуулж, хамгийн ойр сургууль хүртэлх зайг тооцоолсон. Түүнчлэн замын хөдөлгөөний бүсчлэл, инфляц, барилгын зардлын индекс зэрэг өгөгдөл цуглуулсан. Цаашлаад эдгээр өгөгдлүүдийг нэгтгэн цэвэрлэж, 0.8:0.2 харьцаатайгаар сургалтын (training) ба туршилтын (testing) өгөгдлүүдийг хуваан, сургалтын өгөгдлөө ашиглан машин сургалтын алгоритмыг сургаж загвараа үүсгэсэн. Эцэст нь туршилтын өгөгдлийг ашиглан машин сургалтын аргуудыг түгээмэл хэрэглэгддэг үнэлгээний аргуудыг ашиглан харьцуулан үнэлнэ.

Бидний өгөгдлийн багц нь үндсэн хоёр эх сурвалжтай ба онцлог шинж чанар болон түүний эх сурвалжийг Хүснэгт 1-т үзүүлэв.

1. Үл хөдлөх хөрөнгийн мэргэжлийн байгууллага болох "Тэнхлэг зууч"ХХК-аас авсан өгөгдөл;
2. Google maps ашиглан 242 улсын болон 65 хувийн сургуулийн уртраг, өргөргийн өгөгдөл;

<sup>4</sup><https://tosk.gov.mn/>

Үндэсний статистикийн хороо<sup>5</sup>, зам тээврийн яам<sup>6</sup>, үндэсний статистикийн хорооны вэб сайтуудаас цуглуулсан барилгын зардлын индекс, авто замын хөдөлгөөний бүсчлэл (хэвийн, ачаалалтай, их ачаалалтай, маш их ачаалалтай), инфляцийн өгөгдөл;

Бид орон сууцад хамгийн ойр сургуулийн зайг дараах аргаар тооцоолов. 1-р эх сурвалжийн орон сууцныг  $A_i, \overline{1, n}$ , 2-р эх сурвалжийн сургуулиудыг  $B_j, \overline{1, m}$  гэж тэмдэглэв (Зураг 2). Бид эхлээд  $i$  бүрийн хувьд  $A_i$ -с бүх  $B_j$  хүртэлх зайг тооцоолж (тасархай зураас), түүнээс хамгийн бага зайг (ногоон шулуун зураас)

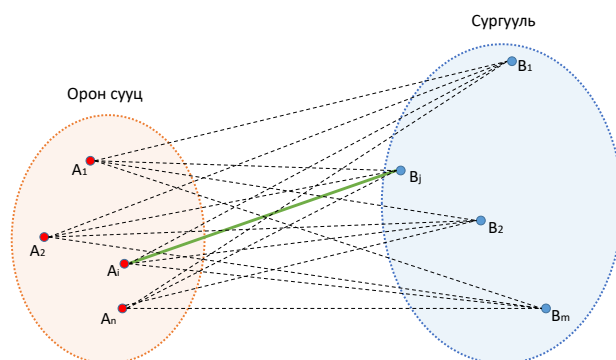
$$\min_j d(A_i, B_j) \quad (1)$$

томъёогоор олсон. Үүнд 2 цэгийн хоорондох зайг

$$d(a, b) = 2r \arcsin \sqrt{\sin^2\left(\frac{\phi_1 - \phi_2}{2}\right) + \cos(\phi_2) \cdot \cos(\phi_1) \cdot \sin^2\left(\frac{\lambda_1 - \lambda_2}{2}\right)} \quad (2)$$

Хаверсины томъёогоор олов. Энд,  $d$  нь зай(км),  $r=6371$ (км),  $\phi$  нь өргөрөг,  $\lambda$  нь уртрагийг илэрхийлнэ.

Бидний өгөгдлийн багц хэр жигд тархалттай байгааг Зураг 3-т харуулав. Зургаас харахад бидний цуглуулсан өгөгдөл нь хотын урд болон төвийн бүсчлэлд илүү тархалттай байна. Цаашлаад сургуулийн зай, орон сууцны үнэ хоёрын хамаарлыг Зураг 4-т харуулав. Үр дүнгээс харахад 1км дотор орон сууцны тархалт өндөр буюу ихэнх орон сууцад ойрхон улсын эсвэл хувийн сургууль байгааг илтгэж байна. Түүнчлэн сургуулийн зай холдох тусам орон сууцны үнэ бага зэрэг буурч байгааг ажиглаж болно.



Зураг 2: Орон сууцтай хамгийн ойр сургууль хүртэлх зайг тооцоолсон байдал. А олонлогийн орон сууц бүрийн хувьд B хүртэлх зайг тасархай зураасаар, хамгийн бага зайг ногоон шулуун зураасаар тэмдэглэв.

<sup>5</sup><https://www.1212.mn/mn>

<sup>6</sup><https://mrttd.gov.mn/>

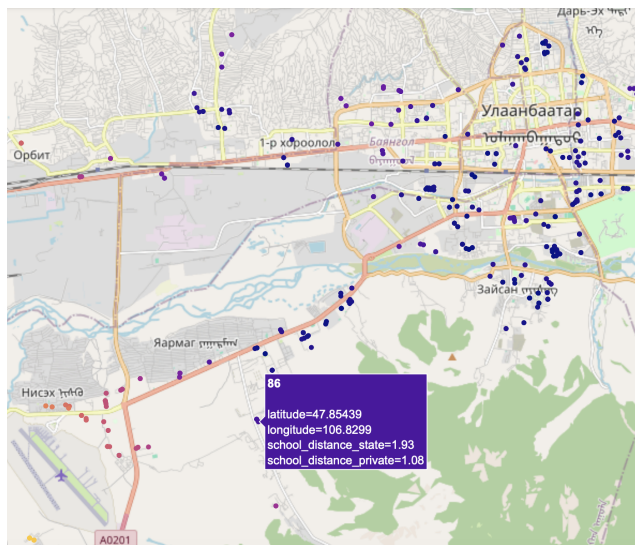
Хүснэгт 1: Орон сууцны үнийг таамаглахад ашиглах онцлог шинж, эх сурвалж

| Онцлог                                         | Төрөл      | Хүрээ             | Мэдээллийн эх сурвалж    |
|------------------------------------------------|------------|-------------------|--------------------------|
| Барилгын төлөв                                 | Хоёртын    | 0,1               |                          |
| Өрөөний сонголт                                | Ангилал    | 0,1,2,3,4,5,6,7,8 |                          |
| Урамшуулал                                     | Ангилал    | 0,1,2,3           |                          |
| Ариун цэврийн өрөөний тоо                      | Ангилал    | 1-4               |                          |
| Барилгын явц                                   | Хоёртын    | 0,1               |                          |
| Хотын төвөөс..км..                             | Тасралтгүй | 0,5-45            |                          |
| Борлуулалтын үйл явц                           | Тасралтгүй | 0-100             |                          |
| Бүсийн нэр                                     | Ангилал    | 0,1,2,3,4,5       |                          |
| Дулаан зогсоолтой эсэх                         | Хоёртын    | 0,1               |                          |
| Дүүргийн нэр                                   | Ангилал    | 0,1,2,3,4,5       |                          |
| Зогсоолын үнэ                                  | Тасралтгүй | 0-45000000        |                          |
| Хороогоор                                      | Тасралтгүй | 1-37              |                          |
| Хамгаалалтын хашаатай эсэх                     | Хоёртын    | 0,1               |                          |
| Өргөрөг                                        | Тасралтгүй | 47,50-48,02       | Тэнхлэг зууч ХХК         |
| Харуул хамгаалалттай эсэх                      | Хоёртын    | 0,1               |                          |
| Уртраг                                         | Тасралтгүй | 106,49-160,60     |                          |
| Тоглоомын талбайтай эсэх                       | Хоёртын    | 0,1               |                          |
| Орон сууцны өрхийн тоо                         | Тасралтгүй | 10-450            |                          |
| Ачааны цахилгаан шаттай эсэх                   | Хоёртын    | 0,1               |                          |
| Барилгын зэрэглэл                              | Ангилал    | 0,1,2,3           |                          |
| Зорчигч цахилгаан шаттай эсэх                  | Хоёртын    | 0,1               |                          |
| Ашиглалтын хугацаа                             | Тасралтгүй | 1-6               |                          |
| Халаалтын нөхцөл                               | Хоёртын    | 0,1               |                          |
| Ашиглалтанд орсон хугацаа                      | Тасралтгүй | 1-100             |                          |
| Барилгын давхарын тоо                          | Тасралтгүй | 2-25              |                          |
| Инфляцийн түвшин                               | Ангилал    | 0,1,2,3,4,5       | Үндэсний статистик хороо |
| Барилгын зардлын индекс                        | Ангилал    | 0,1,2,3           |                          |
| Замын түгжрэлийн бүсчлэл                       | Ангилал    | 0,1,2             | Зам, тээвэр хөгжлийн яам |
| Ойролцоох сургууль хүртэлх зай (км-ээр)        | Тасралтгүй | 0-3, 56           |                          |
| Ойролцоох улсын сургууль хүртэлх зай (км-ээр)  | Тасралтгүй | 0-3, 56           | Google map               |
| Ойролцоох хувийн сургууль хүртэлх зай (км-ээр) | Тасралтгүй | 0, 19-6, 82       |                          |

Бид энэхүү ажлаар орон сууцны үнийг таамаглахад түгээмэл (14; 15; 16) хэрэглэгддэг (хоёрдугаар бүлэгт дурьдагдсан аргууд) машин сургалтын аргууд болох Linear regression (LR), Extreme Gradient Boosting (XGBoost), Support Vector Regression

(SVM), Ridge Regression (RR), Random Forest Regression (RF), Lasso Regression (LaR), ElasticNet (EN), AdaBoost Regression, Decision Tree Regression (DT) аргуудыг ашиглах болно. Ялангуяа Linear, Lasso, Ridge regression аргууд нь өнгөрсөн үеэс өнөөг хүртэл орон сууцны үнэ тогтооход байнга хэрэглэгддэг уламжлалт болон шугаман регрессийн аргууд бол XGBoost, random forest (RF), decision tree (DT) нь сүүлийн үед түгээмэл ашиглагдаж буй илүү үр дүнтэй аргууд юм.

Бүх алгоритмуудын гүйцэтгэлийг тогтвортой байдлын коэффициент  $R^2$ , дундаж абсолют алдаа (MAE), дундаж квадрат алдаа (MSE), язгуур дундаж квадрат алдаа (RMSE) хэмжүүрээр үнэлэв. Эдгээр нь машин сургалт болон статистикт өргөн ашиглагддаг алдааг хэмжих хэмжүүрүүд бөгөөд загварын урьдчилсан таамаглал болон бодит өгөгдлийн хоорондох зөрүүг илэрхийлнэ. MAE нь тухайн загварын урьдчилсан таамаглал болон бодит утгуудын хоорондох абсолют зөрүүгийн дундаж хэмжээг илэрхийлдэг. MSE нь урьдчилсан болон бодит утгуудын хоорондох зөрүүгийн квадратын дундажийг илэрхийлдэг бол RMSE нь MSE-ийн язгуур юм.  $R^2$  нь загварын урьдчилсан таамаглал нь бодит өгөгдлүүдтэй хэр сайн тохирч байгааг илэрхийлдэг.



Зураг 3: Өгөгдлийн тархалт (байршилаар)

Хүснэгт 2: Машин сургалтын арга ашиглан орон сууцны үнийг таамагласан үр дүн

| Арга     | Эхний үр дүн |      |      |      | Эхний үр дүн+нэмэлт |      |      |      |
|----------|--------------|------|------|------|---------------------|------|------|------|
|          | R2           | MAE  | MSE  | RMSE | R2                  | MAE  | MSE  | RMSE |
| LR       | 0.78         | 0.41 | 0.41 | 0.64 | 0.81                | 0.39 | 0.26 | 0.51 |
| XGBoost  | <b>0.87</b>  | 0.24 | 0.23 | 0.48 | <b>0.95</b>         | 0.16 | 0.06 | 0.24 |
| SVM      | 0.73         | 0.36 | 0.50 | 0.71 | 0.87                | 0.26 | 0.17 | 0.42 |
| RR       | 0.78         | 0.41 | 0.41 | 0.64 | 0.81                | 0.39 | 0.26 | 0.51 |
| RF       | <b>0.87</b>  | 0.25 | 0.24 | 0.49 | <b>0.94</b>         | 0.18 | 0.07 | 0.27 |
| LaR      | 0.73         | 0.42 | 0.51 | 0.71 | 0.8                 | 0.38 | 0.28 | 0.53 |
| EN       | 0.78         | 0.41 | 0.41 | 0.64 | 0.81                | 0.39 | 0.26 | 0.51 |
| AdaBoost | 0.83         | 0.37 | 0.32 | 0.57 | 0.88                | 0.31 | 0.16 | 0.4  |
| DT       | 0.82         | 0.34 | 0.35 | 0.59 | 0.86                | 0.29 | 0.18 | 0.43 |

## 4 Туршилт ба үр дүн

Туршилтад бид өмнө дурьдагдсан машин сургалтын аргуудыг scikit-learn (17) санг ашиглан хэрэгжүүлсэн.

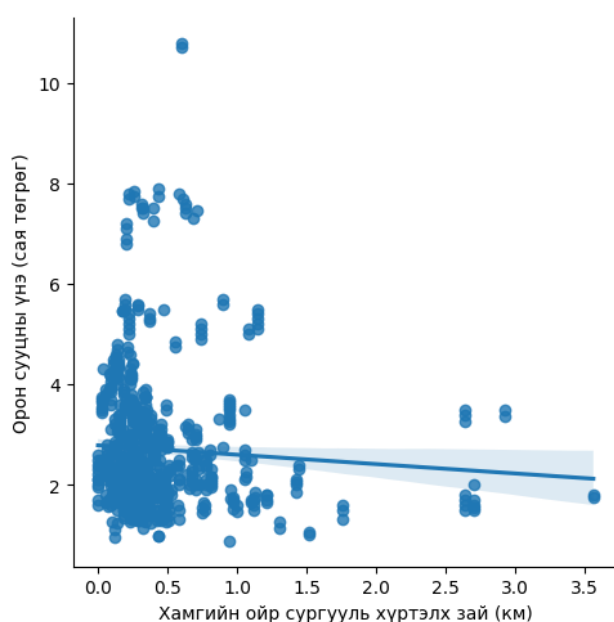
Туршилтын өгөгдөл дээрх машин сургалтын аргуудын үр дүнг Хүснэгт 2-т үзүүлэв. Бүлэг 3-д дурьдсан 1-р эх сурвалжаас авсан өгөгдөл дээр таамагласан үр дүнг “Эхний үр дүн” баганад, 1 болон 2-р эх сурвалжийн өгөгдөл дээр таамагласан үр дүнг “Эхний үр дүн+нэмэлт” баганад харуулав. Үр дүнгээс харахад аль ч машин сургалтын аргууд нь онцлог шинж чанарууд нэмэгдэхэд дундаж квадрат алдаа болон бусад алдааны үзүүлэлт багасч байгааг харж байна. Ялангуяа XGBoost болон RF машин

сургалтын аргууд нь хамгийн сайн үр дүнг үзүүлсэн байна. Тодруулбал, XGBoost загварын  $R^2$  нь 0.87-с 0.95 болж, RMSE алдаа нь 0.48-с 0.24 болж буурсан бол RF загварын  $R^2$  нь 0.87-с 0.94 болж, RMSE нь 0.49-с 0.27 болж буурсан сайн үр дүн үзүүлсэн байна. Бид Хүснэгт 1-т үзүүлсэн онцлог шинжүүд таамаглалын үр дүнд хэрхэн нөлөөлж буй нөлөөллийг тооцоолж үзсэн. Зураг 5-аас харахад орон сууцны үнэд construction rank буюу барилгын зэрэглэл хамгийн их нөлөөтэй байна. Түүнчлэн зогсоолын талбай, өрөөний сонголт болон бидний 2-р эх сурвалжийн өгөгдөл болох сургуулийн зай өндөр хэмжээгээр нөлөөлж буйг харж байна. Тэр дундаа улсын сургуулийн байрлал илүү нөлөөтэй байгааг харж болно.

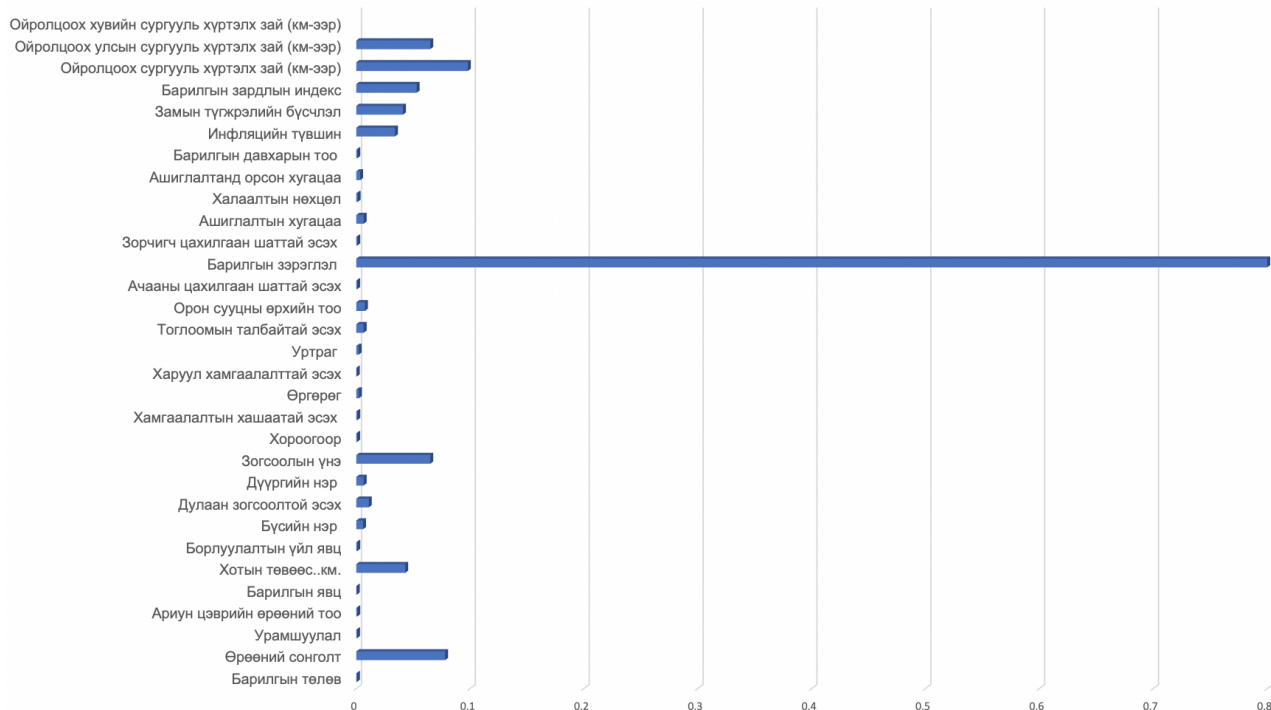
Бодит үнэ болон урьдчилан таамагласан үнэ хоорондын хамаарлыг Зураг 6-т харуулав. XGBoost (зүүн) болон RF (баруун) машин сургалтын аргын таамагласан дүн босоо тэнхлэгийн дагуу бодит үнэ хэвтээ тэнхлэгийн дагуу байх ба үр дүнгээс харахад улаан шугам дагуу тархсан байгаа нь бидний таамагласан үр дүн бодит үр дүнтэй дүйж байгааг харуулж байна.

## 5 Дүгнэлт

Энэхүү ажлаар бид орон сууцны үнийг таамаглахдаа үл хөдлөхийн мэргэжлийн байгууллагаас авсан 25 онцлог шинж чанар бүхий өгөгдөл дээр нэмэлтээр хамгийн ойр сургууль хүртэлх зай, замын хөдөлгөөний бүсчлэл, инфляци, барилгын зардлын индекс зэрэг өгөгдөл цуглуулж машин сургалтын арга ашиглан орон сууцны үнийг таамаглах оролдлогыг амжилттай хийж гүйцэтгэлээ. Үл хөдлөх хөрөнгийн мэргэжлийн байгууллагын өгөгдөл дээр хийсэн үр дүнгээ 80%-95% хүртэл амжилттай ахиулан гайхалтай үр дүнд хүрч чадлаа. Туршилтын үр дүнгээс харахад онцлог шинж чанар нэмэгдэхэд алдааны үзүүлэлт багасч машин сургалтын есөн арга



Зураг 4: Орон сууцны үнэ, сургуулийн зай хоорондын хамаарал.



Зураг 5: Онцлог шинжс чанаруудын нөлөө

нь сайн үр дүн өгснөөс Random Forest болон XGBoost нь сайн хувилбар болохыг баталж чадсан.

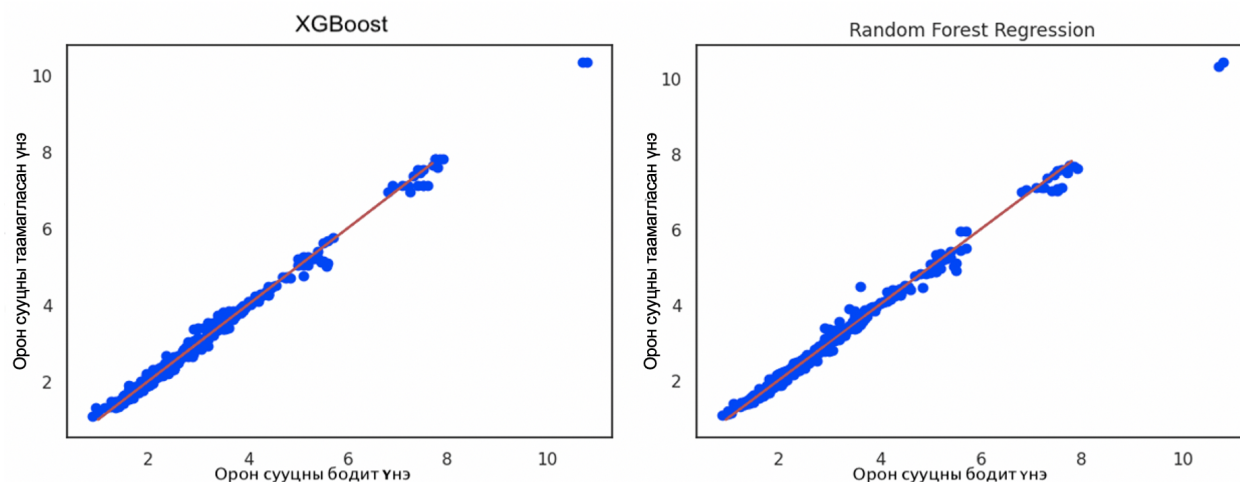
Монгол улсын хувьд орон сууц худалдах, худалдан авахад орон сууцтай ойролцоо улсын болон хувийн сургуулиудын байршил нь орон сууцны үнэд нөлөөлдөг болохыг баталж харуулж чадлаа. Энэ нь Монгол улсын хувьд анх удаа хийгдэж байгаа гэдгээрээ шинэлэг юм.

Цаашдын судалгаанд энэ судалгаанд ашигласан аргуудтай хамгийн сүүлийн үеийн хандлагыг харьцуулах, түүнчлэн судалгааны хүрээнд ашигласан аргуудад илүү өргөн интервалтай гипер параметруудийг турших боломжтой. Нэмж дурдахад, үл

хөдлөхийн үнийг таамаглан үзүүлдэг вэб програм болон аппликейшн хийж мөн газрын зурагт байршуулан орон сууцны үнийг нь ойролцоох сургууль, цэцэрлэг, эмнэлэг, үйлчилгээний төв гэх мэт хэрэгцээт бүхнийг багтаан, хийж хөгжүүлэх боломжтой нь харагдаж байна.

## Зохиогчийн оролцоо

Энэхүү судалгааны ажлын өгөгдөл цуглуулах, цэвэрлэх, машин сургалтын алгоритмуудын туршилтуудыг Ж.Отгончимэг гүйцэтгэж, өгүүллийн бичилтийг Ж.Отгончимэг, Б.Наранчимэг нар хамтран



Зураг 6: Орон сууцны бодит үнэ болон таамагласан үнэ

гүйцэтгэв.

## Ашиг сонирхлын зөрчилгүйн баталгаа

Ашиг сонирхлын зөрчилгүй болно.

## Ашигласан ном

- [1] Fack G, Grenet J. When do better schools raise housing prices? Evidence from Paris public and private schools. *Journal of public Economics*. 2010;94(1-2):59-77. Available from: <https://doi.org/10.1016/j.jpubeco.2009.10.009>.
- [2] Zheng S, Hu W, Wang R. How much is a good school worth in Beijing? Identifying price premium with paired resale and rental data. *The Journal of Real Estate Finance and Economics*. 2016;53:184-99. Available from: <https://doi.org/10.1007/s11146-015-9513-4>.
- [3] Sah V, Conroy SJ, Narwold A. Estimating school proximity effects on housing prices: The importance of robust spatial controls in hedonic estimations. *The Journal of Real Estate Finance and Economics*. 2016;53:50-76. Available from: <https://doi.org/10.1007/s11146-015-9520-5>.
- [4] Yi Lu VS. The Impacts of Public Schools on Housing Prices of Residential Properties: A Case Study of Greater Sydney, Australia. 2023. Available from: <https://doi.org/10.3390/ijgi12070298>.
- [5] Feng Y, Jones K. Comparing multilevel modelling and artificial neural networks in house price prediction,. 2015 2nd IEEE Int Conf Spat Data Min Geogr Knowl Serv. 2015. Available from: <https://doi.org/10.1109/ic SDM.2015.7298035>.
- [6] Отгончимэг Наранчимэг Монгол улсын үл хөдлөх хөрөнгийн үнийг машин сургалтын арга ашиглан таамаглах нь. Монголын мэдээллийн технологи. 2024.
- [7] Varma DSNR Sarma. "House Price Prediction Using Machine Learning and Neural Networks". In 2018 Second International Conference on Inventive Communication and Computational Technologies. 2018;59:1936-9. Available from: <https://doi.org/10.1109/icicct.2018.8473231>.
- [8] Geerts M, De Weerd J, et al. A survey of methods and input data types for house price prediction. *ISPRS International Journal of Geo-Information*. 2023;12(5):200. Available from: <https://doi.org/10.3390/ijgi12050200>.
- [9] Choy LHT. The Use of Machine Learning in Real Estate Research. IEEE Access. 2023;9. Available from: <https://doi.org/10.3390/land12040740>.
- [10] Cihan Cilgin HG. Machine learning methods for prediction real estate sales prices in Turkey. 2023. Available from: <https://doi.org/10.7764/rdlc.22.1.163>.
- [11] Ziming Liu JY. The Effect of School Quality on House Prices: Evidence from Shanghai, China. 2022. Available from: <https://doi.org/10.3390/land11111894>.
- [12] Ruth Ema Febrita HT. Modeling House Price Prediction using Regression Analysis and Particle Swarm Optimization. Malang, East Java, Indonesia. 2017. Available from: <https://doi.org/10.14569/ijacsa.2017.081042>.
- [13] Huang P. Impact of Distance to School on Housing Price. In: Evidence from a Quantile Regression; 2017. .
- [14] Şahinler S. The basic principles of constructing a linear regression model using the least squares method. *Mustafa Kemal Üniversitesi Ziraat Fakültesi Dergisi*. 2000;5(1):57-73.
- [15] Walker JB E Birch. Influence measures in ridge regression. *Technometrics*. 1998;30(2):221-7. Available from: <https://doi.org/10.2307/1270168>.
- [16] Chen C T Guestrin. XGBoost: "A Scalable Tree Boosting System",. In: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining; 2016. p. 785-94. Available from: <https://doi.org/10.1145/2939672.2939785>.
- [17] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. 2011;12:2825-30.