



ИХ ӨГӨГДЛИЙН ТӨВ БАНКАН ДАХЬ ХЭРЭГЛЭЭ

Д.Болдбаатар¹, Б.Дуламзаяа², Б.Эрхэмбаяр³, Ж.Энхбаяр⁴

Хураангуй: Энэхүү судалгааны ажлаар их өгөгдлийн талаарх ерөнхий ойлголтыг танилцуулж, түүнийг хадгалах, хэрэглэх, удирдах тухай, их өгөгдлийг шинжлэх арга техникүүдийн тухай, болон бусад төв банкуудад их өгөгдлийг хэрхэн ашиглаж, үр шимийг нь хүртэж буй туршлагыг түүвэрлэн танилцууллаа. Тоймлон авч үзвэл Төв банкууд их өгөгдлийг эдийн засгийн төлөвийг тодорхойлох, богино хугацааны төсөөлөл хийх, олон нийттэй харилцах харилцааг сайжруулах, санхүүгийн тогтвортой байдал, хяналт шалгалтын чиглэлээр тэргүүн эгнээнд ашиглаж байна. Бусад төв банкуудад их өгөгдлийг ашиглаж буй туршлагаас үзэхэд их өгөгдөл нь Монголбанкны үйл ажиллагааг сайжруулахад чухал хувь нэмэртэй болох нь харагдлаа. Улмаар Монголбанканд их өгөгдлийн удирдлагыг хөгжүүлж, ирээдүйд хэрэглэхэд их өгөгдлийн нэгдсэн сан, түүний дэд бүтцийг тодорхой болгох, хууль эрх зүйн зохицуулалтыг сайжруулах, мэдээллийн аюулгүй байдлыг хангах, их өгөгдлийг шинжлэх арга техникийг судлан, нэвтрүүлэх шаардлага зүй ёсоор тавигдаж байна. Эцэст нь Монголбанк энэ бүх асуудлыг багтаасан их өгөгдлийн удирдлага, зохицуулалтын тухай дунд хугацааны стратегийн баримт бичиг боловсруулан хэрэгжүүлэх нь зүйтэй.

Түлхүүр үгс: *Их өгөгдөл, Төв банк, Машин сургалт, Эх хэлний боловсруулалт, Эдийн засаг, Төсөөлөл, Өгөгдлийн удирдлага*

JEL ангилал: C8

¹ Монголбанкны Мөнгөний бодлогын газрын ЭЗШБХ-ийн ахлах эдийн засагч.

Цахим хаяг: boldbaatar@mongolbank.mn

² Монголбанкны Судалгаа, статистикийн газрын Судалгааны хэлтсийн захирал.

Цахим хаяг: dulamzaya@mongolbank.mn

³ Монголбанкны Мөнгөний бодлогын газрын ЭЗШБХ-ийн эдийн засагч.

Цахим хаяг: erkhembayar.b@mongolbank.mn

⁴ Монголбанкны Мөнгөний бодлогын газрын ЭЗШБХ-ийн эдийн засагч.

Цахим хаяг: enkhbayar.j@mongolbank.mn

I. Удиртгал

Дэлхий даяар цахим хэрэглээ нэмэгдэж, үүнийг дагаад шинэ боловч богино настай мэдээлэл үүсэж, улмаар түүнийг хэрэглэн нийтийн сайн сайханд тустай эсвэл арилжааны өртөг бүтээх аргачлал эрчимтэй хөгжиж, хил хязгааргүй түгэн тархаж байна. Үүний том төлөөлөл болох машин сургалт, гүнзгий сургалтын арга техникууд нь 21-р зууны эхэнд онолын түвшинд байсан бол сүүлийн жилүүдэд тэдгээрийг бодит амьдрал дээр турших, түүний үр шимийг хүртэх боломж нэмэгдсээр байна. Үүнд, тоон өгөгдлөөс гадна бичвэр, дуу, видео бичлэг, зураг зэргийг шинжлэх арга техник улам бүр нарийсан хөгжиж байна. Тухайлбал, бизнесийн сүүлийн үеийн чиг хандлагыг тодорхойлох, эрүүл мэндийн асуудлаас урьдчилан сэргийлэх, гэмт хэрэгтэй тэмцэх зэрэгт олон төрлийн их хэмжээний мэдээллийг ашиглах нь тун үр дүнтэй байна.

Ирвинг Фишерийн хорооны судалгаанд хамрагдсан 52 төв банкны хувьд их өгөгдлийг үйл ажиллагаандаа ашиглаж байгаа төв банкуудын тоо 2015 онд ердөө 30 орчим хувь байсан бол 2020 онд 80 гаруй хувьд хүрч нэмэгдсэн, үүнээс дийлэнх нь дээд удирдлагын түвшинд их өгөгдлийн хэрэглээг нэн чухал асуудал болгон авч үзэж байна (*BIS, 2021*). Их өгөгдлийн тусламжтайгаар төв банкууд эдийн засгийн төлөвийн талаарх мэдээллийг албан ёсны статистик мэдээлэл зарлагдахаас урьдаж, цаг алдалгүй авах боломж нэмэгджээ. Ялангуяа хөгжсөн улсуудын төв банкууд энэ чиглэлд техник технологийн болон хүний нөөцийн чадавхаа дээшлүүлж, тэргүүн эгнээнд алхаж байна.

Одоогоор Монголбанканд их өгөгдлийг бүрдүүлж, түүнд тулгуурлан шинжилгээ хийж эхлээгүй байгаа ч их өгөгдөлтэй холбоотой хэд хэдэн төслүүд Монголбанкны оролцоотойгоор хэрэгжиж эхлээд байна. Цаашид ч Монголбанк хөгжлийн энэхүү чиг хандлагаас хоцролгүй их өгөгдөл цуглуулах, хадгалах, шинжилгээний арга техникийг нэвтрүүлэх шаардлага бий. Тиймээс бид энэхүү судалгаагаар их өгөгдлийн үнэ цэн, түүний төв банканд дахь хэрэглээг бусад улсуудын жишээгээр судалж, Монголбанкны үйл ажиллагаанд их өгөгдлийг нэвтрүүлэх үед анхаарах асуудлууд болон санал, зөвлөмж боловсруулахыг зорилоо. Судалгааны 2-р бүлэгт их өгөгдлийн талаарх ерөнхий ойлголтын талаар, 3-р бүлэгт их өгөгдлийг хэрхэн удирдах талаар, 4-р бүлэгт их өгөгдлийг шинжлэх арга, техникуудийг товч танилцуулав. Харин 5-р бүлэгт их өгөгдлийг ашиглаж буй олон улсын туршлагын талаар (*хавсралтад дэлгэрэнгүй байдлаар*), 6-р бүлэгт Монгол Улсад их өгөгдлийг хөгжүүлэхэд анхаарах асуудал, санал, зөвлөмжийг тусгалаа.

II. Их өгөгдлийн талаарх ойлголт

1. Зарим тодорхойлолтууд

Ерөнхий байдлаар тодорхойлбол “Их өгөгдөл” гэдэг нь уламжлалт арга техникээр шинжилж, үнэ цэнтэй мэдээлэл гарган авахад хүндрэлтэй, зарим тохиолдолд бүр боломжгүй гэж хэлж болох асар их хэмжээний, хурдтай нэмэгдэж байгаа өгөгдлийг хэлнэ. Хэдийгээр их өгөгдлийн талаар бүх нийтээр хүлээн зөвшөөрсөн тодорхойлолт үгүй боловч Doug Laneу-ын 2001 онд дэвшүүлсэн “Гурван V”-н зарчмаас эхлэн тайлбарлавал ойлгоход илүү дөхөм болох юм. Түүний тодорхойлсноор их өгөгдөл нь “volume”, “velocity”, “variety” гэх 3 шинжийг агуулсан байх ёстой байна. Үүнээс гадна сүүлийн жилүүдэд эрдэмтэд, судалгааны байгууллагуудын дэвшүүлээд буй нэмэлт хоёр “V” буюу “veracity”, “value”-г энд нэмж дурдаж болно. Эдгээрийг тус бүрд нь дэлгэрүүлэн авч үзвэл:

“Volume” (хэмжээ) – Технологийн дэвшилтэй уялдан өгөгдлийг цуглуулах, хадгалах зардал үлэмж буурсан, хувь хүний ердийн амьдралын хэв маягаар өгөгдөл бий болж дэлхий дээрх нийт өгөгдлийн хэмжээг нэмэгдүүлэх боломж ихэссэн. Өгөгдлийн хэмжээ маш их байх нь их өгөгдөл гэж нэрийдэх эхний шалгуур болох юм.

“Velocity” (хурд) – Электрон төхөөрөмжийн тоо нэмэгдэхтэй зэрэгцэн бизнесийн ертөнц дэх мэдээллийн урсгал асар хурдасч, түүнийг цаг тухайд нь шинжлэх шаардлага нэмэгдэж байна. Тиймээс өгөгдөл үүсэх, дамжуулах хурд нь их өгөгдлийг илэрхийлэх нэг чухал үзүүлэлт гарцаагүй мөн билээ.

“Variety” (төрөл) – Өгөгдөл ямар ч хэлбэртэй байж болно. Өнгөрсөн хугацаанд цэгцэлж, эрэмбэлсэн бүтэцтэй тоон өгөгдлийг ихээр ашигладаг байсан бол өгөгдлийн шинэ эринд тодорхой бүтэцгүй үг, өгүүлбэр, и-мэйл, видео бичлэг, дуу авиа, санхүүгийн шилжүүлэг гэх зэрэг олон төрлийн мэдээллийг ашиглах болсон.

“Veracity” (үнэн зөв байдал) – Энэ нь өгөгдлийн эх үүсвэр, төрөл, эхний боловсруулалт зэрэгт итгэж болохуйц байх тухай ойлголт юм. Илүү нарийвчлал өндөртэй, итгэл даахуйц өгөгдлийн хэрэгцээ үргэлж байдаг ч илүү хямд, илүү их өгөгдөл цуглуулах тохиолдолд нарийвчлал, найдвартай байдлыг хайхрахгүй орхих явдал элбэг тохиолддог. Тухайн өгөгдөл үнэн зөв байх нь чухал байдлаараа тэргүүн эгнээнд эрэмбэлэгдэнэ.

“Value” (үнэ цэн) – Энэ нь цуглуулсан өгөгдлийг ашиглан үнэ цэн бүтээх тухай ойлголт юм. Их өгөгдлийг үнэ цэн нь байгууллага цуглуулсан өгөгдлөөс өөрийн үйл ажиллагааг сайжруулахад шаардлагатай хэр их мэдээлэл гарган авч чадаж байгаагаас хамаардаг.

Их өгөгдлийн талаарх ерөнхий тодорхойлолтоос гадна дараах ойлголтууд ч мөн дараа дараагийн хэсэгт олон удаа давтагдах тул тус бүрийн тодорхойлолтыг авч үзье.

Бүтэцтэй болон бүтэцгүй өгөгдөл – Тодорхой зорилгын хүрээнд цэгцтэй ангилан бэлдсэн мэдээллийг бүтэцтэй өгөгдөл гэх бөгөөд түүнийг шинжлэх арга техник сайн хөгжсөн байдаг. Харин тодорхой *бүтэцгүй* өгөгдлийг цэгцлэх, хэрэгцээт мэдээллээ гарган авах нь илүү төвөгтэй. Тухайлбал, бүтэцгүй өгөгдлийн жишээ болгож зураг, бичлэг, твийт, бичвэр зэргийг дурдаж болно. Зарим тохиолдолд хагас бүтэцтэй буюу дундын ангилалд хамаарах өгөгдлүүд ч байдаг.

“Internet of things (IoT)” гэдэг нь интернэтэд холбоотой төхөөрөмж, эд зүйлсийг нийтэд нь илэрхийлэх ойлголт юм. Үүнд, ухаалаг утас, ухаалаг цаг, хөргөгч гээд олон биет зүйлсийг багтаан ойлгоно. Маш хямд компьютерын чип, утасгүй сүлжээ өргөн дэлгэрснээр флаш шиг жижиг хэмжээтэй зүйлээс эхлээд онгоц шиг овортой зүйлийг хүртэл IoT-н нэг хэсэг болгох боломжтой болжээ. IoT-н гол давуу тал нь хүний оролцоогүйгээр бодит цагийн горимд мэдээллийг дамжуулахад оршиж байдаг. Эдгээр төхөөрөмжүүд өөрийн ажиллагааны хүрээнд мэдээлэл цуглуулж, түүнийгээ интернэтийн тусламжтайгаар дамжуулснаар **их өгөгдөл** үүсэх нөхцөл бүрддэг.

Hadoop – Энэ нь их хэмжээний өгөгдлийг үр дүнтэйгээр хадгалах, ашиглах боломж олгодог нээлттэй эх¹-ийн технологи юм. Ингэхдээ нэг серверийг олон тооны машинаар өргөтгөж, их хэмжээний өгөгдлийг тэдгээр машинуудад зэрэг байршуулан ачаалдаг. Серверт холбогдсон аль нэг машины үйл ажиллагаа доголдоход түүнийг өөр машинаар орлуулан үйл ажиллагааг хэвийн үргэлжлүүлэх боломжтой байдаг нь энэхүү технологийн найдвартай, хурдан ажиллагааг баталгаажуулдаг.

“SQL” болон “NoSQL” – “SQL (*Structured Query Language*)” нь бүтэцтэй өгөгдлийг хүснэгт (*мөр, баганатай*) хэлбэрээр өгөгдлийн санд хадгалах, сангаас өгөгдөл татан авах үйлдлийг энгийн, үр дүнтэй гүйцэтгэх үндсэн арга юм. Харин “NoSQL (*Not only SQL*)” бүтэцгүй өгөгдөлтэй ажиллахад тохиромжтой юм.

“Data lake” (*Өгөгдлийн нуур*) – Энэ нь нуурт цутгах ус адил олон эх үүсвэрээс хүлээн авсан задгай мэдээллийг анхны форматаар хадгалах сав/нуур юм. Тус газарт орших элемент бүр өөрийн гэсэн таних тэмдэгтэй байдаг. Олон салаагаар нуурт цутгах ус адил мэдээллийг янз бүрийн эх үүсвэрээс авч хадгалж байдаг тул үүнийг өгөгдлийн нуур гэж нэрлэх болжээ.

¹ Хэн ч эх кодыг үзэх, шинэчлэн өөрчлөх боломжтой байдаг программ хангамж юм.

“Data warehouse” (Өгөгдлийн агуулах) – Энэ нь ялгаатай эх үүсвэрээс мэдээллийг цуглуулж, хадгалж түүнээс үнэ цэнтэй мэдээлэл гаргаж өгдөг боловсронгуй электрон агуулах юм. Өөрөөр хэлбэл энэхүү агуулахад өгөгдлийг мэдээлэл болгон хувиргадаг.

“Data lakehouse” (Өгөгдлийн нуурын байшин) – Data lake болон Data warehouse-ын элементүүдийг хооронд нь нэгтгэж зардал багатайгаар хадгалах хэсгийг ийнхүү нэрлэдэг. Энэ нь өгөгдлийн шинжлэх ухаанчдын хувьд машин сургалтын арга техник, бизнесийн ухаалаг систем нэвтрүүлэхэд ихээхэн тус болдог.

“Clouds” (Үүлэн технологи) – Энэ нь дэлхий даяар өгөгдлийн төвүүдэд байршиж, өөр хоорондоо интернэтээр холбогдсон серверийн тухай ойлголт юм. Үүлэн технологийн ачаар хүмүүс ижил файл, программ дээр бараг бүх төрлийн төхөөрөмжүүдээс хандах боломж нээгдсэн.

2. Их өгөгдлийн түүхэн хөгжил

Хэдийгээр их өгөгдөл гэх ойлголт харьцангуй шинэ хэдий ч их хэмжээний өгөгдлийн тухай ойлголт 1960-1970 оны үеэс эхтэй юм. Тухайн үед анхны өгөгдлийн төвүүд, өгөгдлийн бааз хөгжүүлэлт эхэлж байсан түүхтэй. Харин “Их өгөгдөл” гэх нэршлийн хувьд 1990-ээд оны эхээр хэрэглэгдэж эхэлсэн байна. Анх 1989-1990 оны хооронд Тим Бернерс-Ли (*Tim Berners-Lee*), Роберт Кайо (*Robert Cailliau*) нар World Wide Web платформыг бүтээж, улмаар HTML, URL, HTTP-г хөгжүүлсэн бөгөөд интернэтийн эрин үетэй зэрэгцэн 1996 оноос өгөгдлийг цаасаар бус электрон хэлбэрээр хадгалах нь илүү зардал багатай болж, тэрхүү мэдээлэл рүү хандахад ч хялбар болсон байна.

Улмаар 1998 оноос www.google.com домайн үйл ажиллагаагаа эхлүүлсэн бөгөөд тус онд Карл Стрози “NoSQL” өгөгдлийн сангийн хэлийг хөгжүүлсэн түүхтэй. Улмаар 1999 онд Хэл Р.Вэриан (*Hal R. Varian*), Питер Лайман (*Peter Lyman*) нар “How much information” номынхоо анхны хувилбарыг бичиж, тус номдоо дэлхий дээрх нийт дижитал мэдээллийн хэмжээг тодорхойлохыг оролдож байв.

Түүний дараагаар 2000 оноос хойш интернэт болон вебийн тусламжтайгаар өмнө нь байгаагүй их хэмжээний мэдээллийг цуглуулах, шинжлэх боломж нээгджээ. Тухайлбал, вебийн ачаалал, онлайн дэлгүүрийн тоо нэмэгдсэний дээр Yahoo, Amazon, eBay зэрэг томоохон компаниуд үйл ажиллагаа эхэлж, цаашлаад хэрэглэгчдийнхээ зан төлөвийг даралтын хувь², байршил, хайлтын түүх зэргээс нь шинжиж мэдэхээр ажилласан. Энэ нь цаашлаад өгөгдлийн шинэ ертөнцийг бидэнд нээн өгч байна.

² *click-through rate*

Их өгөгдлийн өсөлт хөгжил нь “Hadoop” зэрэг нээлттэй эхийн технологийн хөгжилтэй салшгүй холбоотой юм. Facebook, YouTube болон бусад онлайн платформ өөрсдийн хэрэглэгчдийн талаар маш их хэмжээний өгөгдөл үйлдвэрлэж байгаа нь олны анхаарлыг татах болов. Улмаар 2005 онд “Hadoop” технологийн хөгжүүлэлт дуусаж, олон нийтэд танилцуулагдаж, “NoSQL” хэл нэрд гарч эхлэв. Эдгээр технологийн тусламжтайгаар өгөгдлийг хямд зардлаар хадгалах боломжтой болсон. Улмаар сүүлийн хэдэн жилд их өгөгдлийн хэмжээ асар хурдтай нэмэгдсэн бөгөөд хэрэглэгчид одоо ч их хэмжээний өгөгдлийг тасралтгүй үйлдвэрлэж байна.

Компьютерын хүчин чадал хурдтай нэмэгдэж буй энэ үед “их өгөгдөл” гэх ойлголтыг тоон утгаар илэрхийлж, хязгаар тогтоох боломжгүй юм. Анх 1965 онд Гордон Мүүр (*Gordon Moore*) нэг чипт багтаах транзиторын тоо буюу компьютерын хүчин чадал 2 жил тутамд 2 дахин нэмэгдэх таамаглал дэвшүүлсэн нь биеллээ олж хожмоо Мүүрийн хууль хэмээн алдаршсан байдаг. Сүүлийн 60 орчим жилийн тооцоогоор компьютерын хүчин чадал 18 сар тутамд 2 дахин нэмэгджээ. Гэсэн хэдий ч Гордон Мүүрийн үзэж буйгаар нэг чипт багтаах транзиторын тоог хязгааргүй ихэсгэх боломжгүй бөгөөд компьютерын хүчин чадлын өсөлтийн хурд цаашид саарч болзошгүй байна.

Өдгөө өгөгдөл хөрөнгийн нэг хэлбэр болж хувирч байна. Дэлхийн томоохон технологийн компаниудын хувьд тэдний бүрдүүлсэн өгөгдөл нь байгууллагын үнэ цэнийг тодорхойлох томоохон хүчин зүйл болохын зэрэгцээ шинэ бүтээгдэхүүн боловсруулах гол орц болж байна. Өгөгдлийн хэмжээ нэмэгдэж, хадгалах зардал буурахын хэрээр оновчтой бизнесийн эсвэл бодлогын шийдвэрт ашиглаж байна.

Их өгөгдлийн үнэ цэн нь танд хэр их өгөгдөл байгаагаас бус байгаа өгөгдлийг хэрхэн ашиглаж байгаагаас хамаарна. Өгөгдлийг цэгцлэх, шинжилгээнд тохируулж бэлдэх нь ихээхэн цаг шаардах бөгөөд өгөгдлийг шинжилдэг хүмүүс нийт ажлын цагийнхаа 50-аас 80 хувийг зөвхөн өгөгдлийг цэгцэлж бэлдэхэд зарцуулдаг гэх тооцоо бий. Олон улсын өгөгдлийн корпорацаас (*IDC*) гаргасан тооцоогоор дэлхий нийтээр бизнесийн зорилгоор цуглуулсан бүтэцгүй өгөгдлийн 90 хувьд огт шинжилгээ хийж чадахгүй байна. Шинэ технологи гарч ирснээр өгөгдлийг хадгалах боломж нэмэгдэж байгаа ч өгөгдлийн хэмжээ хоёр жил тутамд хоёр дахин нэмэгдэж байгаа гэдгийг санах нь зүйтэй юм. Тиймээс байгууллагууд тэр их мэдээллийг зохицуулах үр дүнтэй арга хайсан хэвээр л байна.

3. *Их өгөгдөлтэй холбоотой статистик*

- 2022 оны байдлаар олон улсын их өгөгдлийн салбар 274.3 тэрбум ам.доллараар үнэлэгдэж байна.

- Google хуудсанд өдөрт дунджаар 3.5 тэрбум хайлт бүртгэгддэг.
- WhatsApp аппликэйшнээр өдөрт дунджаар 100 тэрбум давсан зурвас илгээгддэг.
- Дэлхийн нийт бизнесүүдийн 95 хувь нь бүтэцгүй өгөгдлийг удирдаж чадахгүй байна.
- 2021 онд 79 зеттабайт³ өгөгдөл үйлдвэрлэгдсэн бол энэ тоо 2025 онд 180 зеттабайт болж нэмэгдэхээр байна.
- Интернет хэрэглэгчид өдөрт 2.5 их ингүүмэл (10^{18}) байт мэдээлэл үйлдвэрлэдэг.
- 2020 оны статистикаар интернет хэрэглэгчид дунджаар 1.7 МБ/сек хурдаар мэдээлэл үйлдвэрлэж байна.
- Өнөөдрийн байдлаар нийт байгууллагуудын 91 хувь нь хиймэл оюун ухаан болон их өгөгдөлд тодорхой хэмжээгээр хөрөнгө оруулсан байна.
- Үйлдвэрлэгдэж буй нийт өгөгдлийн 80-аас дээш хувь нь бүтэцгүй өгөгдөл байна.
- Өгөгдөлтэй холбоотой хийгдэж буй үйлдлийн тоо (үүсгэх, хэрэглэх, хуулах гэх мэт) өмнөх 10 жилд 5000 хувиар нэмэгдсэн байна.
- Өндөр хурдны интернет ашиглан өнөөдөр интернетэд хадгалагдаж буй нийт өгөгдлийг татаж авахад 181 сая жил шаардагдана.
- Дэлхий даяар интернет хэрэглэгч 4.3 тэрбум хүн өдөрт дунджаар 6 цаг 42 минутыг интернетэд зарцуулж байна. Энэ нь нийлбэртээ нэг өдөрт бид 1.2 тэрбум жилийг онлайн орчинд өнгөрүүлж, мэдээлэл үйлдвэрлэж байна.
- Компаниуд их өгөгдлийн үнэ цэнийг сүүлийн жилүүдэд маш ихээр ойлгож эхлээд байна. Тухайлбал, Fortune 1000 компаниудын 62 хувь нь аль хэдийн өгөгдөл хариуцсан захиралтай болоод байна. 2012 онд энэ тоо ердөө 12 хувийн орчимд байв.
- Банкны сектор дахь их өгөгдлийн зах зээл 2025 он гэхэд 62 тэрбум ам.долларт хүрэхээр байна. Энэ үзүүлэлт 2012 онд 12 тэрбум ам.доллар, 2019 онд 30 тэрбум ам.доллар байсан бол 2025 он хүртэл жилд дунджаар 13 хувиар өсөх хүлээлттэй байна.
- Хэрэглэгчдийн онлайн орчинд зарцуулж буй нийт хугацааны 33 хувийг Social media бүрдүүлдэг.
- Фейсбүүкийг сар бүр идэвхтэй ашигладаг нийт 2.8 тэрбум хэрэглэгч байдаг.

³ 1 зеттабайт = 10^{21} байт

- Твиттерийн хэрэглэгчид 1 минутад дунджаар хагас сая гаруй удаа жиргэдэг.
- Дэлхий дээр 1 секунд бүрд 127 шинэ төхөөрөмж интернэтэд холбогдож байна. Энэ хурдаар үргэлжилбэл 2025 он гэхэд IoT төхөөрөмжийн тоо 41.6 тэрбум ширхэгт хүрэх юм.

III. Их өгөгдлийг удирдах нь

1. Их өгөгдлийн удирдлагын талаарх ерөнхий ойлголт

Интернэтийн орчинд амьдралын хэрэгцээгээр хэрхэн мэдээлэл үүсэж, тэрхүү мэдээлэлд маш их үнэ цэн агуулагдаж байдаг тухай өмнөх бүлэгт тайлбарласан. Харин үнэ цэнийг бодит болгоход чухал нөлөөтэй зүйл нь **өгөгдлийн удирдлага**⁴ буюу тухайн өгөгдлийг ашиглах, хандах, хамгаалах, арчлах тухай ойлголт юм. Өгөгдлийн сайн удирдлага нь тухайн байгууллагын шийдвэр оновчтой байх, дүрэм журмын хэрэгжилтийг хангах зэрэгт эерэг нөлөөтэй.

Мэдээллийг шинжлэх, хэрэглэх, хадгалах технологи үсрэнгүй хөгжих тусам хувь хүний болон хуулийн этгээдийн мэдээллийг ашиглан нэмэгдсэн өртөг бүтээх боломж тэлж байна. Өмнө нь мэдээлэлтэй хамаатай хүн, хуулийн этгээдийг тодорхойлох боломжгүй байсан бол технологийн хөгжлөөс шалтгаалж мэдээллээс нь тухайн хүнийг тодорхойлох боломж улам бүр нэмэгдэж байна. Түүнчлэн, мэдээллийг хэрэглэгчид тохируулах арга хөгжих тусам хүний мэдэх эрхийг технологиор хязгаарлах боломж ч бий боллоо.

Асар өндөр хурдацтай нэмэгдэж буй их өгөгдлийн том хэсэг нь “хүнийг тодорхойлох боломжтой”⁵ мэдээлэл бөгөөд хэдийгээр эзэмшигч нь мэдээлэл ашиглах зөвшөөрөл өгсөн боловч тэр нь хаашаа, хэнд дамжиж байгааг хянах боломжгүй болжээ. Өгөгдлийн сан эзэмшиж байгаа этгээд дээрх аргыг ашиглан өгөгдлийн шинжилгээнээс арилжааны үнэ цэн, нийгмийн амьжиргаанд эерэг үр нөлөөтэй сурвалж бий болгож чаддаг болсон учир хүн, нийгмийн бүлэг, хуулийн этгээдийн эрхийг хязгаарлахгүй, гадуурхахгүй, хүний хувийн болон эмзэг мэдээлэл, цахим тодорхойлогчид халдахгүй байх шаардлага үүсэж байна. Үүнтэй холбоотой эрх зүй үүсэж, нийгмийн харилцааны шинэ *орон зай* бүрдэж эхэллээ. Манай улсын хувьд Хүний хувийн мэдээллийг хамгаалах тухай, Нийтийн мэдээллийн ил тод байдлын тухай Монгол улсын хуулиудыг 2021 оны 12 сард баталж 2022 оны 5 сараас мөрдөж эхэлж, эрх зүйн хэм хэмжээ тогтоосон билээ.

⁴ Data governance

⁵ Хүний хувийн мэдээллийг хамгаалах тухай Монгол улсын хуульд тодорхойлсон утгаар хэрэглэв.

Их өгөгдөлтэй холбоотой нийтлэг асуудлыг авч үзэхэд улс орны хил, эрх зүйн орчныг тооцож үзэх нь хангалтгүй юм. Энд яригдаж буй өгөгдөл виртуал орчинд хил хязгааргүй үүсэж, хадгалагдаж байна. Их өгөгдлийн шинжилгээний хэрэгслүүд нь мөн виртуал орчинд даяаршин түглээ. Үүний сайн тал нь олон улсын шилдэг туршлага, технологийг хялбар, шуурхай хуваалцах боломжтой болжээ. Гэхдээ манай мэдээллийг бид ашиглаж виртуал орон зайгаа эзэмших, манай иргэдийн сайн сайхны төлөө ашиглахын төлөө хил хязгааргүй орчинд өрсөлдөх сорилт тулгаж байна. Өөрөөр хэлбэл, даяар орчлонгийн их өгөгдлийг ашиглан өртөг бүтээх үндэсний тогтолцоотой болох нь эдийн засгийн ач холбогдолтой болжээ. Их өгөгдөл өөрөө нэмүү өртөг бүтээх өндөр өгөөжтэй салбар болж буй учраас дотооддоо их өгөгдөл хэрэглэх боломжийг нээж өгөх тусам бид гадаадын бизнест төлбөр төлөх нь багасаж, бизнесийн орчингоо Монголдоо үлдээх юм. Үүний тулд өгөгдлийн бодлого, цахим орчны ёс зүй, нийгмийн сэтгэл зүй нь бизнесийн таатай нөхцөл бүрдүүлэхүйц байх хэрэгтэй болно. Өгөгдлийн бодлогоо тодорхойлж, хэрэгжүүлэхэд энэ эрх зүйн орон зай чухал боловч нийтээр хүлээн зөвшөөрсөн ёс зүйн хэв шинж бий болгох үүнээс чухал хэрэгтэй байгааг 21 дүгээр зууны өгөгдлийн бодлогын тухай санал солилцсон мэргэжилтнүүд зөвлөж байна.

Их Британийн Эзэн хааны нийгэмлэг Британийн Академитай хамтран 2018 онд “Өгөгдлийн удирдлага, хэрэглээ” төсөл хэрэгжүүлж, түүвэр судалгаа хийж, семинар зохион байгуулсан бөгөөд төслийн тайланд ёс зүйн хэв шинжид сорилт учруулж буй асуудлуудыг дэвшүүлжээ. Эдгээрийн заримаас дурдвал:

1. Мэдээллийг нийтийн ашиг сонирхлын эсвэл арилжааны зорилгоор хэрэглэхдээ түүний олдоч, сонголтыг хязгаарлахгүй байх тэнцвэрийг олох,
2. Боловсруулалт, шинжилгээний үр өгөөжийг нийтэд ил хэрэглэхдээ хүн, нийгмийн бүлэг, хуулийн этгээдийн ашиг сонирхолд харшлах түвшинг зохистой хэмжээнд хязгаарлах,
3. Өдөр тутмын амьдрал, хүн хоорондын харилцаанаас үүсэж байгаа өгөгдлийг ашиглан өртөг бүтээхдээ “хүний хувийн мэдээлэл”⁶, эрхийг хүндэтгэх; инновацыг урамшуулахдаа нийтийн ашиг сонирхлын зорилготой байлгах тэнцвэрийг хангах,
4. Хүн, хуулийн этгээдийн мэдээллийг хамгаалахдаа хувь хүн, нийтийн эрх ашгийн төлөө мэдээлэл хуваалцах, хэрэглэх, шинжлэх *орон зайг хангахаар* зохион байгуулах буюу өгөгдлийг шинэ аргаар боловсруулахыг сэдэлжүүлэхдээ уг мэдээлэл нь нийтийн эрх ашгийн төлөө дамжих, зарагдах боломжтой байхыг чухалчлах; мэдээлэл дамжиж байгаа үйл ажиллагааны ил тод байдлыг хангах ёс зүйтэй байх зэрэг асуудлуудыг тусгасан байна.

⁶ Хүний хувийн мэдээллийг хамгаалах тухай Монгол улсын хуульд тодорхойлсон

Эдгээр анхаарууштай асуудлууд нь орчин үеийн их өгөгдөл боловсруулах технологи ашиглахдаа өмнө нь хүнийг тодорхойлох боломжгүй байсан мэдээлэл цахим орчинд хүн, хуулийн этгээдэд хамааруулах боломж үүсэх, түүнд нөлөөлөх чадвартай болбол дээрх тэнцвэрийг хангах нийгмийн ойлголцолд хүрэх сорилт тулгаж байна. Түүнчлэн, хүнд шууд мэдрэгдэхгүй ч хүний эрхийг хязгаарлах тохиолдол үүсэж болно. Жишээ нь: хиймэл оюун ухаан ашиглан хүний сонголтыг үндэслэн “тохиромжтой” мэдээллээр хангах зорилт тавих үед хүний мэдэх эрхийг хязгаарладаг бөгөөд энэ үед “тохируулах”, “хязгаарлах” хоёрыг хүлээн зөвшөөрөхүйцээр дэнслэх хэрэгтэй гэжээ.

2. Монголбанк ба Их өгөгдөл

Монголбанк цаашид их өгөгдлийг бодлогын шийдвэрт ашиглах зорилгоор нийтийн сайн сайхны төлөө хэрэглэх шаардлага үүсэх нь зүй. Үүнийг маргаангүй, зорилгын дагуу цуглуулах, боловсруулах, хэрэглэх, нийтийн мэдээллийн нээлтэй өгөгдөл болгоход “Цахим хөгжлийн багц хууль” болон “Нийтийн мэдээллийн ил тод байдлын тухай” Монгол Улсын хуулиудыг мөрдлөг болгоно. Эдгээрийн хуулиудад тодорхойлсон хэм хэмжээ нь хүний хувийн болон эмзэг мэдээлэл, цахим тодорхойлогчийг нийтийн мэдээлэл болгохдоо мэдээллийн эзний эрх ашгийг зөрчихгүйгээр хүн, хуулийн этгээд болон нууцлалыг тодорхойлох боломжгүй болгон боловсруулж, нээлтэй өгөгдөл болгож болно⁷ гэжээ. Нөгөөтээгүүр, хүний хувийн мэдээллийг хамгаалах хуульд Төрийн байгууллага “мэдээллийн эзний зөвшөөрлөөр”⁸ мэдээлэл цуглуулах, боловсруулах, ашиглахаар зохицуулсан байна. Эдгээр зохицуулалтууд нь их өгөгдлийн хувьд хоорондоо зөрчилтэй хэрэгжих, өртөг бүтээх боломж нь хязгаарлагдмал байх юм. Үүнээс гадна “мэдээлэл цуглуулах, боловсруулах, ашиглахад баримтлах зарчим”⁹ бүгд хамгаалах зохицуулалт бөгөөд Монголбанк шиг хуульд зөвшөөрсөн зохицуулалтын дагуу ажиллах байгууллагыг их өгөгдөл ашиглах боломж муутай болгож байна. Мэдээллийн “бүрэн бүтэн байдлыг алдагдуулахгүй байх” нь боловсруулалтын зарим аргачлалд зохистой эсэх нь эргэлзээтэй байна. Хүний мэдээллийг хэт хамгаалах нь өгөгдөл цуглуулж, шинжилж, хэрэглэх, нийтэд ил болгоход саад учруулах магадлалтай байвал их өгөгдлөөр бодлогын шинжилгээнд хэрэглэх үр дүн гаргаж авч, нийтийн сайн сайхны төлөө хэрэглэх боломжгүй болно. Энэ утгаараа өгөгдлөөс бий болгох өртөг алдагдах юм. Түүнчлэн, Google мэтийн хил хязгааргүй орчинд бий болж буй өгөгдлийг Монгол улсаас гадна боловсруулж өртөг бүтээх боломжтой болж, бид цахим орон зайгаа эзэмших бололцоо хумигдана.

⁷ *Нийтийн мэдээллийн ил тод байдлын тухай Монгол улсын хуулийн 12.1 дүгээр заалт*

⁸ *Уг хуулийн 6.1.1, 6.2.1 дүгээр заалтууд*

⁹ *Уг хуулийн 5 дугаар зүйл*

Монголбанк их өгөгдөл ашиглах нь хүний нөөцийн, технологийн шинэ шийдэл олох ажил болох юм. Төв банк статистикийн хуулийн дагуу боловсруулж, нийтэд ил болгодог мэдээллээс гадна төлбөрийн тооцоотой холбоотой өндөр давтамжтай өгөгдөл ялгаж авах боломжтой бөгөөд их өгөгдлийг ашиглахад хэрэглэх мэдээллийг бусад мэдээлэл хуримтлуулагчдаас дамжуулан эсвэл худалдаж авах боломжтой. Дэлхий даяар төв банкууд их өгөгдлийн аргачлалаар “Тухайн улирлын төсөөлөл”¹⁰ хийж байна. Тэдний өгөгдөл цуглуулах, ашиглах, хэрэглэх, хадгалахад тавигдаж буй нийтлэг шаардлагын үндэс нь хүн, хуулийн эрх ашгийг хамгаалах бол нөгөөтээгүүр өгөгдлийг ёс зүйтэй хэрэглэх юм¹¹. Хэрэглээний ёс зүйг хянах, зөвшөөрөхөд туслах үүднээс “Өгөгдлийн санч”¹²-тай байх шаардлагатай. Өгөгдөлтэй ажилладаг байгууллагууд дотоод нэгжүүддээ санд хандах, дамжуулах үүрэгтэй ажлын байр бий болгодог байна. Өгөгдөл өндөр давтамжтай байхаас гадна олон янз, жишээ нь: зарим нь тоон, зарим нь үгээр илэрхийлэгдсэн байх учир түүнийг шинжлэх аргууд нь мөн олон янз байх болно. Зарим аргачлал нь уламжлалт тоон шинжилгээний аргачлалаас хэтэрхий өөр учраас түүнийг боловсруулахад тусгай мэргэшил шаардлагатай. Тиймээс ч өгөгдлийн санчдын сүлжээ, тэдний хамтын ажиллагаа нь шинжилгээний арга хуваалцах, өртөг бүтээх процессын нэг хэсэг болж метадатаг мөрдөж өгөгдлийг шинэчлэх үүрэг гүйцэтгэдэг байна.

Бусад төв банкуудын өгөгдлийн удирдлагын тогтолцоог нэгтгэж үзэхэд хоёр хэлбэр давамгайлж байна. АНУ, ИБУИНВУ, Сингапурын төв банкууд тусдаа өгөгдлийн асуудал эрхэлсэн удирдлага бий болгон ажиллагааг нь интеграцид оруулж байгаа бол ОУТТБ, Бундесбанк, АНУ-ын төв банкны Нью Йоркийн банк зэрэг нь өгөгдөл боловсруулах тусгай төв үүсгэн ажиллаж байна. Аль ч хэлбэрээр ажиллалаа гэсэн цахим өгөгдлийн шинэчлэгдэж буй орчин, уламжлалт бус шинэ аргачлал, бусад орнуудад хэрэглэгдэж эхлээд буй аргачлал нь хиймэл оюун ухаан, машин сургалт зэрэг технологийн, хүний нөөцийн инновац шаардсан үргэлжилсэн процесс болж байна. Өгөгдлийн хэлбэр олон янз байх тусам түүнийг макро эдийн засаг, санхүүгийн зах зээлийн шинжилгээний үнэ бүхий актив болгон хувиргах арга барил олширч, өгөгдлийн санг удирдаж байгаа тогтолцооны нягт, тасралтгүй хамтын ажиллагааг шаардаж байна. Энэ процессын цөмд *мэдээлэл хуваалцах ёс зүй* нэн чухал болжээ. Төв банкн дээр өгөгдлийн удирдлагыг хэрхэн зохион байгуулсан жишээг БНСУ-ын төв банкны туршлагаас Монголбанкны тус байгууллагатай хамтран гаргасан тайлангаас харж болно.

¹⁰ Nowcasting

¹¹ Central Banking Publication, 2016. “Big Data in Central Banks” мэдээллийг нэгтгэсэн байна.

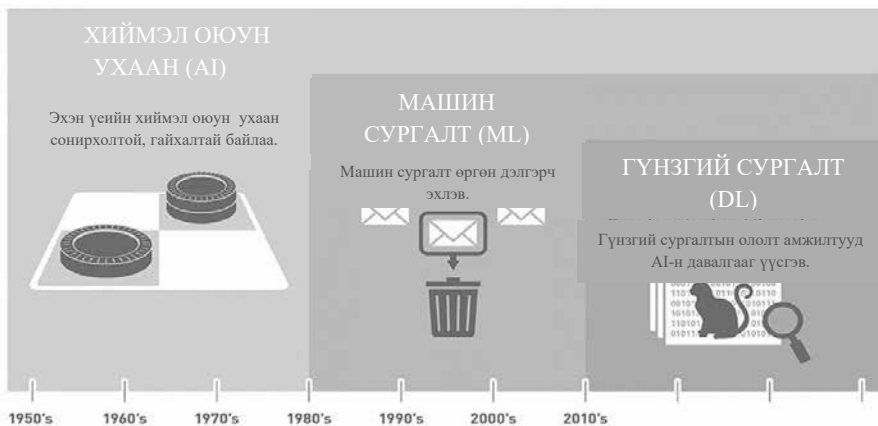
¹² Data steward

Эцэст нь өгөгдлийг шинжлэх, хэрэглэх, хадгалах нь хүний сайн сайхны төлөө хийгдэх учиртай. Ингэхдээ хүний болон нийтийн эрх ашиг, сонирхлыг хүндэтгэх; өгөгдлийг удирдах болон хэрэглэх нь ил тод, хариуцлагатай, хүний эсвэл нийтийн оролцоог хангах; ардчилсан зарчмыг баримталсан байх зарчмуудыг хэрэгжүүлсэн байх шалгуур тавигдаж байх нь зохимжтой. Хүний эрхийн түгээмэл тунхаглалын зарчим ч ийм шалгууртай нийцтэй байх болно. Өгөгдөлтэй харьцах энэ зарчим нь алдаагүй, төгс ажиллагааны баталгаа биш учраас амжилт, алдаанаас суралцах нийтийн ёс зүйтэй, зохистой ойлгуулж чаддаг байхаар өгөгдлийн удирдлагын зохион байгуулах хэрэгтэй болно.

3. Их өгөгдлийг шинжлэх арга, техникүүд

Их өгөгдлийн хэмжээний хувьд их, давтамж өндөр, ялгаатай хэлбэртэй зэрэг шинж чанаруудаас нь шалтгаалан уламжлалт статистик, эконометрикийн арга, техникүүдээр шинжлэхэд маш их хэмжээний нөөц, хүчин чадал, цаг хугацаа зарцуулах шаардлагатай болдог. Харин хиймэл оюун ухаан, түүний дэд хэсгүүдийн тусламжтайгаар их өгөгдлийг шинжлэхэд илүү хялбар болж, түүнийг бүх түвшний бодлого, шийдвэр гаргалтад ашиглах боломжтой болсон. Их өгөгдлийг шинжлэхэд хиймэл оюун ухаан, түүн дотроо машин сургалт, гүнзгий сургалт, эх хэлний боловсруулалт гэсэн арга техникүүдийг түлхүү ашиглаж байна. Эдгээр аргуудын талаар дараагийн хэсэгт нэг бүрчлэн авч үзье.

Зураг 1. Их өгөгдлийг шинжлэх арга техникүүд



Эх сурвалж: Michael Copeland, Nvidia Corporation

3.1 Хиймэл оюун ухаан

Хиймэл оюун ухаан (*artificial intelligence эсвэл AI*) нь хүний сэтгэхүйг дуурайлган, бараг ижил түвшинд хүний үйлдэл, үйл ажиллагааг гүйцэтгүүлэхээр хөгжүүлсэн компьютерын программ хангамж юм. Хиймэл оюун ухааныг ашигласнаар

их өгөгдлийг бэлтгэх ажлыг илүү боловсронгуй, автоматжуулж, өгөгдлийг дүрслэх, таамаглан загварчлах, бусад нарийн төвөгтэй аналитик шинжилгээний даалгавруудыг гүйцэтгэдэг тул өгөгдлийн шинжилгээг илүү хялбар болгодог.

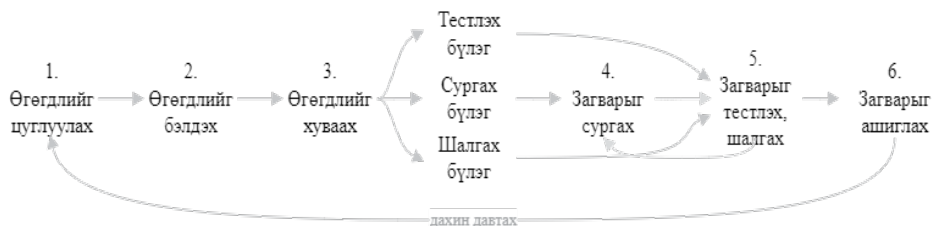
3.2 Машин сургалт

Машин сургалт нь өгөгдсөн ажил эсвэл мэдээллийг тусгай зааварчилгаа буюу программ ашиглалгүйгээр тодорхой алгоритмууд, статистик загварууд ашиглан тухайн өгөгдөл ажлын нийтлэг шинж төлөв дээр үндэслэн судалж, сурч авах үйл ажиллагаа юм. Энэ нь хиймэл оюун ухааны салбарын нэгэн дэд хэсэг болно. Машин сургалтын зорилго нь шинэ өгөгдөлд байнга дасан зохицож, тэдгээрээс шинэ хэв маяг, дүрмийг олж илрүүлэх явдал юм. Заримдаа үүнийг хүний удирдамж, тодорхой программчлалгүйгээр хэрэгжүүлэх боломжтой байдаг. Тиймээс машин сургалтыг ерөнхийд нь удирдлагатай, удирдлагагүй гэж хоёр ангилдаг. Машин сургалт нь сүүлийн үеийн онол, технологийн ололт амжилтаас үүдэн өнөө үед өгөгдлийн шинжлэх ухааны хамгийн хүчтэй хөгжиж буй салбар болжээ.

Машин сургалт нь ердийн программчлагдсан алгоритмаас ялгарах гол ялгаа нь тодорхой программын тусламжгүйгээр өгөгдлийг боловсруулах чадвар юм. Өөрөөр хэлбэл, машин сургалтын хувьд инженерээс өгөгдлийн бичлэгийн төрөл тус бүрийг хэрхэн боловсруулах талаар машинд нарийн зааварчилгаа өгөх шаардлагагүй байдаг. Машин сургалт оролтын өгөгдөлд тулгуурлан эдгээр дүрмийг өөрөө тодорхойлдог.

Аливаа машин сургалтын гүйцэтгэлийн гол олдвор нь түүхэн өгөгдлийн дэд багцаар сургагдсаны дараа алгоритм шинэ өгөгдлийг хэрхэн боловсруулдаг болохыг дүрсэлсэн математик загвар юм. Сургалтын зорилго нь зорилтот утга (*шинж чанар*), өгөгдлийн объект бүрийн үл мэдэгдэх утгыг томъёолох чадвартай загварыг боловсруулахад оршино. Таамаглалын загварыг сургахын тулд та алгоритмд шаардлагатай ялгаатай өгөгдлүүдийг оруулах хэрэгтэй (*үйлчлүүлэгчид орхисон эсвэл худалдан авалт хийсэн эсэхээс үл хамааран*). Энэхүү түүхэн өгөгдлөөс сурч мэдсэнээр загвар нь ирээдүйн өгөгдөлд таамаглал дэвшүүлэх боломжтой болно.

Зураг 2. Машин сургалтын ажлын зураглал



Эх сурвалж: www.centropolis.mn

Ерөнхий дараах алхмуудын дагуу ажилладаг. Үүнд:

Өгөгдөл цуглуулах. Дижитал дэд бүтэц болон бусад эх сурвалжаа ашиглан аль болох их, хэрэгцээтэй өгөгдөл цуглуулж, мэдээллийн санд нэгтгэх.

Өгөгдөл бэлдэх. Өгөгдлийг урьдчилан боловсруулах, цэвэрлэх нь нэлээд боловсронгуй байж болох боловч баганад ижил утгыг өөр хэлбэрээр оруулсныг ижил болгох, дутуу утгыг бөглөх, өгөгдлийн бусад алдааг арилгахыг зорьдог. Жишээлбэл, 2016 оны 12-р сарын 14, 2016.12.14 гэж бичигдсэн бол ижил форматад оруулах шаардлагатай.

Өгөгдлийг хуваах. Загварыг сургахын тулд өгөгдлийн дэд хэсгээр салгаж, сургаж шинэ өгөгдөл оруулахад загвар хэр сайн эсвэл муу ажиллаж байгааг мэдэхэд хэрэглэгдэнэ.

Загварыг сургах. Алгоритм нь өгөгдлийн хэв шинж, зан төлөвийг таньж мэдэх.

Загварыг тестлэх болон шалгах. Өгөгдлийн тестлэх, шалгах дэд бүлгүүдийг ашиглан загварын гүйцэтгэлийг үнэлэх, таамаглал хэр үнэн зөв болохыг ойлгох, тестлэн баталгаажуулах.

Загварыг өргөтгөх. Туршилт хийсэн загварыг аналитик шийдлийн нэг хэсэг болгон шийдвэр гаргах тогтолцоондоо оруулах эсвэл чадварыг хэрэглээнд ашиглах боломжийг олгох (*жишээлбэл, бүтээгдэхүүн санал болгох системийг сайжруулах*).

Дахин сэргээх. Загварыг шат дараалалтай сайжруулахын тулд шинэ өгөгдлийг цуглуулах буюу шинэ өгөгдөл гарах үед процессыг эхнээс нь дахин давтана.

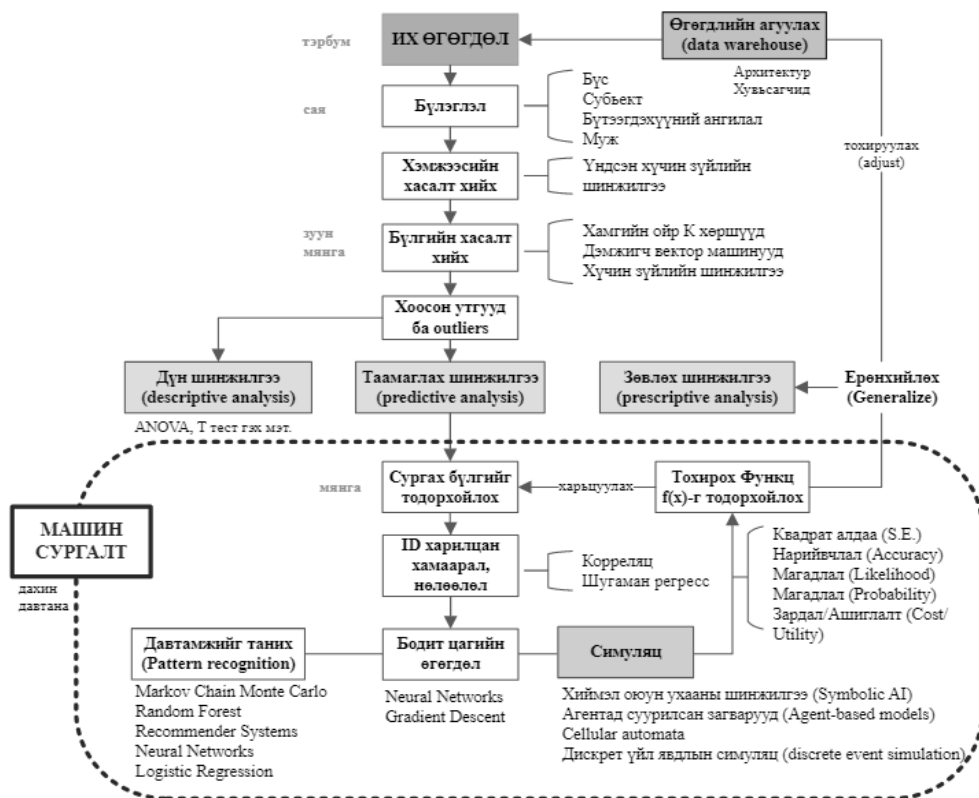
3.3 Машин сургалт ба Их өгөгдөл

Их өгөгдлийг машин сургалттай хослуулснаар хоёр талдаа үр өгөөжтэй. Нэг талаас, олон төрөл зүйлийн, их хэмжээтэйгээр тасралтгүйгээр бий болж байдаг их өгөгдлийг машин сургалтын алгоритмуудад орц болгон ашигласнаар сургалтын өгөгдлийн хэмжээ нэмэгдэхийн хэрээр машин сургалтын алгоритмууд илүү сайжирч, хөгждөг. Нөгөөтээгүүр, машин сургалтын алгоритмууд нь тэрхүү тасралтгүйгээр үүсэж буй их өгөгдлийг **цаг алдалгүйгээр** автоматаар бэлтгэн боловсруулж, хадгалах, өгөгдлөөс нийтлэг хэв шинж, дүрмийг тодорхойлох, аналитик шинжилгээ хийхэд тусалдаг юм.

Их өгөгдөл болон машин сургалтын харилцан уялдааг Зураг 3-г харууллаа. Их өгөгдлийг шинжлэхдээ юуны өмнө их өгөгдлийг бүс, субъект, бүтээгдэхүүний ангилал, муж зэргээс хамааруулан бүлэглэдэг. Үүний дараа үндсэн хүчин зүйлийн шинжилгээ ашиглан хэмжээсний (*dimensional*) хасалт, бүлгийн хасалтыг хийнэ. Улмаар өгөгдөлд дүн шинжилгээ хийж, утга оруулаагүй болон хэт өндөр болон бага

(outliers) утгуудыг хасна. Үүний дараа өгөгдөлд таамаглах шинжилгээ (predictive analysis)-г хийх бөгөөд энэ хэсгээс машин сургалтын хэрэглээ орж ирнэ. Таамаглах шинжилгээний явцад өгөгдлөөс сургах бүлгийг тодорхойлж, өгөгдлийн харилцан хамаарал, нөлөөлөл, давтамж (pattern)-д тулгуурлан загварыг шалгах бөгөөд хиймэл оюун ухаанд суурилсан машин сургалтын олон удаагийн симуляцын дүнд тохирох загвар, функцийг тодорхойлон, уг процессыг маш олон удаа давтдаг. Ийнхүү машин сургалтаар тохирох загвар, функцийг тодорхойлсны дараа түүнийг хувийн секторт болон төрийн бодлогод тооцоо, шинжилгээг баяжуулах, сайжруулах, чанарын өгөгдлөөс нэмэлт мэдээлэл гаргах, зах зээлийн тренд, хэв шинжийг тодорхойлох, зээлийн үнэлгээ хийх, хяналт шалгалт хийх зэрэг төрөл бүрийн шинжилгээнд ашигладаг.

Зураг 3. Машин сургалтыг их өгөгдөлд ашиглах зураглал



Эх сурвалж: Rubens Zimbres

3.4 Гүнзгий сургалт

Гүнзгий сургалт (deep learning) нь хүний мэдлэгт тулгуурлан асуудлыг шийддэг аргуудын нэг дэд бүлэг юм. Энэ сургалт нь өгсөн өгөгдлийн онцлогт суурилж өөрөө сурахын зэрэгцээ өгөгдлийг ангилах, регресс хийх даалгавруудыг давхар

гүйцэтгэж, өгөгдлөөс мэдээлэл гаргаж авдаг. Гүнзгий сургалт нь олон давхаргат/үет шугаман бус фильтерүүдийн системийн тусламжтай өгөгдлийн шинж чанарыг (*feature*) гарган авдаг. Давхарга болгоны оролт нь өмнөх давхаргын гаралтуудтай холбогдсон байна. Тиймээс давхарга болгон дээр тухайн өгөгдлийн параметр, онцлогийг гаргаж авахад чиглэдэг бөгөөд өмнөх онцлогуудаасаа уламжлагдсан мэт байдаг.

Гүнзгий сургалтын архитектурыг биологийн хувьд хүний тархины бүтцээс, ялангуяа харааны кортексээс санаа авч бүтээсэн. Үүнд, өгөгдлөөс зүй тогтол, мэдээллийг ойлгож авахын тулд давхаргууд эсвэл мэдрэлийн эсүүд шугаман бус функцуудыг ашиглан өгөгдлийг ээлж дараалан боловсруулдаг. Эдгээр мэдрэлийн сүлжээний давхарга бүрд багцалсан зургуудыг боловсруулах бөгөөд хамгийн сүүлийн давхарга нь шийдвэр гаргахад шаардлагатай бүх өгөгдлийн онцлог шинжүүдэд хандах боломжтой байдаг. Мэдрэлийн сүлжээ (*neural network*)-ний давхаргын бүтцийг ихэвчлэн ашиглаж байгаа өгөгдлийн төрөлд тохируулан нарийвчилж өгдөг. Гүнзгий сургалтын гол хязгаарлалт нь суралцах олон тооны жишээ хэрэг болдог явдал юм. Учир нь оновчтой ангилагч (*classifier*)-ийг олохоос гадна түүнд тохирох шинж чанаруудыг бас олохын тулд том түүвэр хэрэг болдог.

Гүнзгий сургалт нь машин орчуулга, зураг дүрс, дуу таних, автомат тоглоомын технологи зэрэгт үсрэнгүй өөрчлөлт авчирсан. Харин төв банкуудын хувьд гүнзгий сургалтыг ерөнхий тэнцвэрийн загварын шийдийг олоход түлхүү ашиглаж байна. Тухайлбал, ерөнхий тэнцвэрийн загвар дахь шугаман бус Тейлорын дүрмийн хувьд хоёр тэнцвэрт төлөвийг тодорхойлох, ялгаатай бодлогын дэглэмүүдийн хувьд төсөв, мөнгөний бодлогын уялдааг шинжлэх, загварын агентуудын рациональ хүлээлтийн таамаглалыг сулруулж, агентын хүлээлт суралцдаг байдлаар загварчлах зэргийг дурдаж болно. Гүнзгий сургалтыг ийнхүү шийд олох хэрэгсэл болгон ашиглаж байгаа нь хэд хэдэн шалтгаантай. (i) Онолын хувьд гүнзгий сургалт нь универсал шугаман бус *approximator* юм. Тиймээс олон тэнцвэрт төлөвийг тодорхойлоход ашиглагддаг, (ii) Гүнзгий сургалт нь “*dimensionality curse*”-г шийдэхэд тусалдаг, (iii) Гүнзгий сургалтын алгоритм нь кодлоход хялбар, тогтвортой, их хэмжээний параллелчлалаар дамжуулан өргөтгөх боломжтой, хамгийн сүүлийн үеийн техник, технологийг ашиглан хэрэгжүүлэх боломжтой байдаг.

3.5 Эх хэлний боловсруулалт

Эх хэлний боловсруулалт (*NLP-Natural Language Processing*) нь машин сургалтын нийтлэг хэрэглээний багц бөгөөд бичвэрээс мэдээлэл гаргаж авахад ашиглагддаг. NLP-г мэдрэмжийн шинжилгээ эсвэл машин орчуулга зэрэг удирдлагатай сургалтын асуудлыг шийдэхэд ашиглаж байна. Жишээлбэл, бичвэрийг эерэг

эсвэл сөрөг утгаар нь ангилах, эсвэл хоёр хэлний бичвэрийг харьцуулан тулгаж орчуулга хийх гэх мэт. Харин удирдлагагүй сургалтын хувьд сонины нийтлэл, шинжлэх ухааны нийтлэл гэх мэт баримт бичгийн том цуглуулгад дүн шинжилгээ хийж, түүнийг дараа нь хүмүүс тайлбарлах боломж бүхий сэдвийн загварчлалд ашигладаг. Машин сургалтын NLP-д үзүүлж буй гол нөлөө нь үгсийг тодорхой орон зайд дүрсэлж, үг хоорондын Евклидийн зай нь үгийн утга, үгс хоорондын хамаарлыг илэрхийлэх байдлаар хэлний загваруудыг үүсгэж, хөгжүүлдэг явдал юм.

IV. Их өгөгдлийг Төв банканд ашиглах нь: Олон улсын туршлага

Сүүлийн арванд их өгөгдөл асар хурдтай үүсэж, хуримтлагдан, түүнийг боловсруулах, шинжлэх технологийн дэвшлүүд гарснаар эдийн засгийн бүхий л салбарт их өгөгдлийг ашиглах болсон. Үүний дотор төв банкууд их өгөгдөлд тулгуурласан шинжилгээг бодлого боловсруулахдаа илүү өргөнөөр хэрэглэх болжээ. Төв банкууд их өгөгдлийг мөнгөний бодлогын шинжилгээ, үүн дотроо эдийн засгийн төлөвийг тодорхойлох, макро эдийн засгийн загварчлалд тусгаж, төсөөлөлд ашиглах, олон нийттэй харилцах, санхүүгийн тогтвортой байдлын шинжилгээ зэрэгт түлхүү ашиглаж байна. Эдгээртэй холбоотой судалгааны ажлуудын тоймыг дараах хэсэгт тайлбарлав (*Дэлгэрэнгүй хүснэгтийг хавсралтаас харна уу*).

1. Эдийн засгийн төлөвийг тодорхойлох, төсөөлөл боловсруулах

Сүүлийн жилүүдэд шинжээчид өдрийн давтамжтай, бүтэцтэй болон бүтэцгүй тоон өгөгдлийг эдийн засгийн идэвхжил хэмжих, таамаглал боловсруулахад ихээхэн туршиж байна. Тус төрлийн судалгаанд түгээмэл ашиглагдаж буй өгөгдөлд эдийн засгийн мэдээ, интернэт хайлтын өгөгдөл, төлбөр тооцооны бүртгэл зэрэг багтаж байна.

Нэгдүгээрт, эдийн засгийн мэдээ буюу тухайн мэдээнд агуулагдаж буй үгсийн сентимент шинжилгээг ихэвчлэн тухайн улс орны эдийн засгийн идэвхжилийг өдөр, 7 хоногийн давтамжтай тооцоход ашиглаж байна. Тухайлбал, Испани (*Aguilera нар, 2021*), Бельги (*Algaba нар, 2021*), Итали (*Aprigliano нар, 2021*) зэрэг улсад хийгдсэн судалгаануудаас үзэхэд судлаачид тухайн орны хамгийн их уншигчтай сэтгүүлүүдийн эдийн засгийн сэдэвтэй мэдээг цуглуулж, машин сургалтын NLP, ВМА зэрэг аргачлал ашиглан их өгөгдлөөс өндөр давтамжтай үзүүлэлт ялган авчээ. Эдгээр үзүүлэлтүүд нь эдийн засгийн идэвхжилийн хугарлын цэгийг тодорхойлох, богино хугацааны төсөөлөл боловсруулах чадвар өндөртэй гарсан байна.

Хоёрдугаарт, төв банкууд өөрсдийн төлбөрийн систем дээр хадгалагдаж буй

төлбөр тооцоо, картын гүйлгээний их өгөгдөлд шинжилгээ хийх замаар өрхийн хэрэглээ, эдийн засгийн үйл ажиллагааны “тухайн улирлын төсөөлөл”-ийг хийж байна. Тухайлбал, Испани (*Gil нар, 2018*), Норвеги (*Aastveit нар, 2020*), Канад (*Charman болон Desai, 2021*), Итали (*Aprigliano нар, 2019*), Голланд (*Verbaan нар, 2017*), Энэтхэг (*Rooj болон Sengupta, 2021*)-ийн төв банкууд төлбөр тооцоо, картын гүйлгээний өндөр давтамжтай, бодит цагийн их өгөгдлөөр MIDAS загварыг өргөтгөснөөр эдийн засгийн өсөлт, өрхийн хэрэглээний тухайн улирлын таамаглал боловсруулах чадвар сайжирсан гэх дүгнэлтэд хүрчээ.

Гуравдугаарт, судлаачид иргэдийн гар утасны байршил, интернэт хайлтын үр дүнгээр тэдний үйл хөдлөлийг хэмжиж, цаашлаад эдийн засгийн зарим салбарын өнөөгийн нөхцөл, ойрын төлөвийг тодорхойлох боломжийг эрэлхийлж байна. Тухайлбал, Япон (*Matsumura нар, 2021*) болон Латин Америк, Карибын тэнгисийн орнууд (*Sampi болон Jooste, 2020*)-д хэрэглэгчдийн гар утасны байршлын мэдээлэл ашиглан жижиглэн худалдаа, зочлох болон аж үйлдвэрийн салбарын идэвхжилийг хэмжихийг оролджээ. Судалгааны үр дүнгээр хүн амын хөдөлгөөнт байдлыг илэрхийлэх энэхүү үзүүлэлт нь эдийн засгийн тодорхой салбарын идэвхжилийг бодит цагийн горимоор тодорхойлохоос гадна богино хугацааны төсөөлөл гаргах чадвар өндөртэй байна гэж үзжээ. Мөн D’Amuri болон Marcucci (2017) нар гүүгл хайлтын мэдээлэлд тулгуурлан АНУ-ын ажилгүйдлийн таамаглал боловсруулах загвар хөгжүүлсэн. Тэд тухайн өдрийн нийт гүүгл хайлтад эзлэх “jobs” хайлтын харьцаагаар үзүүлэлт үүсгэж, энгийн авторегрессив загвар ашигласан нь бусад түрүүлэгч индикатортай харьцуулахад таамаглал боловсруулах, эргэлтийн цэгийг тодорхойлох чадвар өндөртэй байжээ. Түүнчлэн, Woloszko (2020) эдийн засгийн хэд хэдэн нэршил/хэллэгийн хувьд “Google Trends” мэдээлэл ашиглан 7 хоногийн давтамжтай “Google Trends Tracker (GTT)” үзүүлэлт боловсруулж, түүнийгээ 46 улсын эдийн засгийн өсөлтийн төсөөлөл боловсруулахад туршсан. Судалгааны үр дүнгээр “GTT” үзүүлэлт нь тэдгээр 46 улсын ихэнхэд нь эдийн засгийн өсөлт, хямрал, бизнесийн мөчлөгийг дохиологч мэдээлэл олгож байна гэж үзжээ.

2. Олон нийттэй харилцахад ашиглах

Шинжээчид машин сургалтын “neural network”, “k-nearest neighbor”, “random forest” аргууд, энгийн регресс болон бусад аргачлал ашиглан төв банкуудын олон нийттэй харилцах харилцааны нийтлэг байдал, үр ашгийг хэмжихийг зорьж байна. Тухайлбал, Gorodnichenko нар (2021) Fed-н ерөнхийлөгчдийн хэвлэлийн хурлын асуулт хариултын үеэрх дууны өнгийг шинжилж, үр нөлөөг тооцохыг зорьсон. Судалгааны үр дүнгээс үзэхэд хэвлэлийн хурлын үеэр “эерэг” дууны өнгөөр олон нийттэй харилцахад хөрөнгийн зах зээлд эерэг нөлөө үзүүлж, ам.долларын ханш

еврогийн эсрэг чангарах нөлөөтэй байдаг гэх дүгнэлтэд хүрсэн байна. Харин Ehrmann болон Wabitsch (2021) нарын хувьд Европын Төв Банктай холбоотой 5.5 сая гаруй Твиттер жиргээг цуглуулж, шинжээчид болон иргэд Европын Төв банкнаас гаргаж буй мэдээлэлд хэрхэн хариу үйлдэл үзүүлж буйг шинжлэхийг зорьжээ. Тэд субъектив шинжтэй үг хэллэг, утга агуулсан мэдээлэлд иргэд илүү хариу үйлдэл үзүүлж, дахин жиргэх магадлал өндөр бөгөөд тухайн мэдээлэл харьцангуй удаан хугацаанд олны анхааралд өртдөг болохыг тогтоожээ. Мөн Baumgartner болон Zahner (2021) нар 1980-2021 оны хооронд BIS-д хэвлэгдсэн болон төв банкуудын сайтад нийтлэгдсэн мэдээлэлд үндэслэн төв банкуудын олон нийтийн харилцааны нийтлэг байдал, үр нөлөөг хэмжсэн байна. Тус судалгааны хувьд (i) хямралын үед ECB зэрэг төв банкуудын гаргаж буй мэдээлэл, тайлангууд нь санхүүгийн зах зээлд оролцогчдын үйл хөдлөлийг намжаах нөлөөтэй, (ii) инфляцыг онилох бодлого хэрэгжүүлэгч төв банкуудаас гаргаж буй мэдээлэл нь агуулгын хувьд FED, ECB зэрэг төв банкуудтай төстэй байдаг гэх дүгнэлт гарчээ.

3. Санхүүгийн тогтвортой байдлын шинжилгээнд ашиглах

Санхүүгийн тогтвортой байдлын шинжилгээнд их өгөгдлийг өрхийн өрийн дарамтыг үнэлэх, зээлийн мэдээллийн сан үүсгэх, зээл олголтыг хянах, банкны балансын үзүүлэлтүүдийн богино хугацааны төсөөлөл боловсруулах, системийн эрсдэлийн урьдчилан дохиологчийг шалгах зэрэгт түгээмэл ашиглаж байна (*Cœuré, 2017*). Үүнээс гадна олон нийтийн мэдээллийн хэрэгсэл, гар утасны хэрэглээ, хэрэглэгчийн төлбөр тооцооны түүхэн цувааг санхүүгийн зах зээлийн хөгжлийг хэмжих, зээлийн өсөлтөөс үүсэх эрсдэлийг тодорхойлох, зорилтот бүлгийн зээлдэгчдийн санхүүгийн чадамжийг үнэлэх, эх үүсвэрийн зардлыг бууруулах зэрэгт ашиглаж байна (*Tian нар, 2018*).

Урьдчилан дохиологч: Машин сургалтын техник ашиглан хямралыг дохиолох үзүүлэлт хөгжүүлэх оролдлого олон улсын хэмжээнд түгээмэл байна. Жишээлбэл, Huang (2019) нар “Financial Times” сэтгүүлийн сүүлийн 40 жилийн мэдээлэл ашиглан сентимент үзүүлэлт тооцож, хямралыг дохиолох чадварыг шалгажээ. Судалгааны үр дүнгээс тус сентимент үзүүлэлт нь санхүүгийн хямралын өмнө (*Бразил 1999*), эдийн засгийн хүндрэлийн үеэр (*Турк 2018*) амжилттай дохиолж байна гэж дүгнэжээ. Мөн сентимент үзүүлэлтүүд¹³-ийн дохиолох чадвар улс орнуудын хувьд ялгаатай байгаа нь ажиглагдсан. Тухайлбал, Бразил улсад “fear” болон “hedging”, Мексик болон Швед улсад “crisis”, Аргентин болон Боливи улсад “opinion” сентиментүүд сайн дохиологч болж байжээ. Мөн Kiley болон

¹³ Сентимент үзүүлэлтүүдийг “fear”, “crisis”, “opinion”, “hedging”, “risk” гэсэн 4 үг ашиглан тооцсон.

D. Vrontos нар (2020) сарын давтамжтай хувьсагчид ашиглан АНУ-ын эдийн засгийн хямралыг таамаглах загварчлал хөгжүүлсэн нь Logit/Probit зэрэг эконометрик загваруудаас таамаглалын чадвар сайн байжээ.

Хяналт шалгалт: Санхүүгийн байгууллагууд өдрөөс өдөрт нэмэгдэж буй асар их өгөгдөлтэй ажиллахдаа олон хууль, дүрэм журам, зохицуулалтыг дагаж мөрдөх, харилцагчийн гүйлгээний нууцлалыг хянах, луйвар болон хууль бус үйл ажиллагаанд хяналт тавих зэрэг сорилтуудтай тулгардаг. Энэхүү асуудлыг шийдэхэд “regulatory technology (*RegTech*)”, “supervisory technology (*SupTech*)”-ийн шийдлүүд тусалж байна. Тухайлбал, эдгээр төрлийн технологи нь дүрэм журмын өөрчлөлтийн түүхийг бүртгэх, технологи ашиглан хяналт тавих, луйврыг илрүүлэх, харилцагчийг таньж мэдэх, терроризмын санхүүжилтийг таслан зогсоох, зохистой эрсдэлийн менежментийг хэрэгжүүлэх, аудитад тохиромжтойгоор тайлан бэлтгэх зэрэгт өндөр үр өгөөжтэй байж, зардлыг хэмнэж байна. Broeders болон Prenio нар (2018) “SupTech”-ийг санхүүгийн хяналт, шалгалтад ашиглах талаар судалж, “SupTech” нь банкуудын мэдээллийн систем дэх задгай их өгөгдөлд шууд хандаж, автоматаар өгөгдлийг нэгтгэх, баталгаажуулах, шинжилгээ хийх замаар макро болон микро зохистой бодлогын хүрээнд хяналт, шалгалтыг дэмждэг бөгөөд SupTech-г ашигласнаар зардал буурч, үр өгөөж нэмэгдэх, хүчин чадал тэлэх давуу талуудтай гэж үзжээ.

Зээлийн мэдээллийн сан: Европын төв банк бусад Төв банкуудтай хамтран “дэлхийн зээлийн бүртгэл” бүхий Анакредит нэгдсэн санг боловсруулж, зээлийн эрсдэлээс сэргийлэхэд ашиглаж байна (*Schubert, 2018*). Уг сан нь Европын бүх банкуудын мэдээллийг нэгтгэж байсан бол одоо цар хүрээгээ улам тэлж байгаа юм. Анакредит нь зээлийн болон зээлийн өсөлтөөс үүсэх эрсдэлийг илэрхийлэгч үзүүлэлтүүдээс бүрддэг микро мэдээллийн сан бөгөөд энэ нь бодлого боловсруулагчид зээлийн өсөлтийг хянах, банкны зээлийн өсөлтийг хязгаарлах чиглэлийн макро зохистой бодлогын шийдвэрүүдийн үр дүнг үнэлэхэд дэмжлэг үзүүлж байна.

V. Дүгнэлт, анхаарах асуудлууд

Бусад улсуудын төв банкуудад их өгөгдлийг ашиглаж буй туршлагаас үзэхэд их өгөгдөл нь төв банкны мөнгөний бодлого болон макро зохистой бодлого боловсруулалт, хяналт шалгалт, олон нийттэй харилцах зэрэг олон талын үйл ажиллагааг сайжруулахад чухал хувь нэмэртэй болох нь харагдлаа. Цаашид Монголбанк их өгөгдлийг ашиглахад нэгдсэн өгөгдлийн сан бий болгох, түүнийг олж авах, ашиглахтай холбоотой хуулийн зохицуулалтуудыг сайжруулах, тодорхой болгох, их өгөгдлийг боловсруулах, удирдах техникийн болон хүний нөөцийн хүчин чадлыг хөгжүүлэх, кибер аюулгүй байдлыг хангах, өгөгдлийн шинжилгээ хийх хиймэл оюун ухааны алгоритмыг хөгжүүлэх зэрэг сорилтууд нэн тэргүүнд тулгарахаар байна. Эдгээрийг тус бүрд нь авч үзье.

Их өгөгдөлтэй холбоотой хууль, эрх зүйн зохицуулалтыг сайжруулах, тодорхой болгох: Хууль эрх зүйн зохицуулалт талаас бусад орны сайн жишээг гүнзгийрүүлэн судалж, их өгөгдлийг ашиглахад шаардлагатай хууль эрх зүйн өөрчлөлтийн саналыг холбогдох байгууллагуудад хүргүүлэх, хамтарч ажиллах хэрэгтэй байна. Үүнд, Монголбанк их өгөгдөл бүрдүүлэх, “төрийн их өгөгдөл”¹⁴-д хандах, хувийн сектороос их өгөгдөл худалдан авах зэрэг харилцааг зохицуулахад чиглэсэн эрх зүйн зохицуулалтын талаар нэн тэргүүнд авч үзэх шаардлагатай байна. Өгөгдлийг ашиглан өртөг бүтээх боломжийн нээж өгөх, ёс зүйтэй ашиглах хязгаарлалтыг зохистой хэмжээнд хэрэглэх зэрэгт анхаарах шаардлагатай.

Их өгөгдлийн сан, түүний дэд бүтцийн архитектурыг тодорхой болгох: Банкуудын системийг Монголбанкны серверт холбох, өгөгдлийн шинжилгээ, дүрслэлийн программ хангамж худалдан авах ажил хэрэгжиж байна. Цаашид технологи, хүний нөөцөд хөрөнгө оруулах замаар энэ ажлыг эрчимжүүлж, банкны салбарын их өгөгдөл бүрдүүлэх, бодит цагийн өгөгдлийг боловсруулдаг болох хэрэгтэй байна. Түүнчлэн, төрийн их өгөгдлийн санд хандаж, банкнаас бусад секторын мэдээллийг хуулийн хүрээнд авах нь Монголбанкны мэдээллийн баазыг сайжруулахад асар том нөлөөтэй байх болно. Зарим төрлийн мэдээллийг Монголбанк дангаар эсвэл Төрийн их өгөгдлийн сангаас гарган авах боломжгүй тохиолдолд хувийн сектортой хамтарч их өгөгдөл бүрдүүлэх, хувийн байгууллагаас их өгөгдөл худалдан авах шаардлага үүснэ.

Мэдээллийн кибер аюулгүй байдал болон бизнесийн тасралтгүй үйл ажиллагааг хангах: Их өгөгдлийн бааз үүсгэж, түүнийг удирдахад мэдээллийн кибер аюулгүй байдлын менежментийн тогтолцоог олон улсын жишигт

¹⁴ Одоогоор хэрэгжилтийн шатандаа явж буй төсөл

нийцүүлэн хөгжүүлэх, түүнийг цахим хөгжлийн багц хуультай уялдуулан олон улсын стандартад нийцүүлэх шаардлага зүй ёсоор үүснэ. Иймд төв банк “Цахим хөгжлийн багц хууль” болон олон улсын хөгжлийн чиг хандлагад нийцүүлэн өөрийн мэдээллийн аюулгүй байдлын менежментийн тогтолцоог MNS ISO/IEC 27001:2021 стандартын дагуу шинэчлэх ажлыг үе шаттайгаар эхлүүлэх, кибер аюулгүй байдлыг хангах, хөндлөнгийн аудитаар үнэлгээ тогтмол хийлгэн, өөрийн мэдээллийн аюулгүй байдлыг хангах тогтолцоог хөгжүүлэх хэрэгтэй. Түүнчлэн, байгалийн гамшиг тохиолдох үед бизнесийн тасралтгүй үйл ажиллагааг хэрхэн хангаж ажиллах талаар төлөвлөгөө боловсруулж ажиллах шаардлага тулгарна.

Их өгөгдлийг шинжлэх арга техникийг судлах, нэвтрүүлэх: Их өгөгдөлд шинжилгээ хийхээр сургасан машин сургалтын алгоритмыг бусад төв банкуудаас судалж, суралцах, нэвтрүүлэх нь хэрэгцээ үүснэ. Түүнчлэн, дунд хугацаанд “өгөгдлийн шинжлэх ухаан”¹⁵, “өгөгдлийн инженер”¹⁶, мэдээллийн аюулгүй байдлын чиглэлээр мэргэжилтнүүд бэлтгэх, бусад ажилчдыг сургах, улмаар дотооддоо машин сургалтын болон бусад сүүлийн үеийн арга, техникийг хөгжүүлэх шаардлагатай байна.

Мэргэжилтнүүд бэлтгэх,

- Төв банк өгөгдлийн шинжлэх ухаан, өгөгдлийн инженер, кибер аюулгүй байдал, өгөгдлийн аюулгүй байдал, өгөгдлийн удирдлагын чиглэлээр мэргэжилтнүүд бэлтгэх, олж авах,
- Ажилчдыг өгөгдлийн шинжлэх ухааны чиглэлээр сурахыг дэмжих, гадаад харилцааны хүрээнд энэ чиглэлийн мэргэжилтнүүдийг илүү дэмжих талаар зөвшилцөлд хүрэх хэрэгцээ шаардлага гарна.

Мэргэжилтнүүдийн ур чадварыг дээшлүүлэх чиглэлээр олон улсын байгууллагуудтай хамтран ажиллаж, сайн туршлагаудыг судлах,

- Одоогийн Монголбанкны ажилчдын судалгаа шинжилгээндээ ашиглаж буй эконометрикийн шинжилгээ хийх зориулалттай программууд нь их өгөгдөлтэй ажиллахад тохиромжгүй тул “Python”, “R”, “Talend”, “Cloudera”, “Tableau” зэрэг программуудыг сурах, шинжилгээ хийх арга техникийг эзэмшихэд гаднын байгууллагуудтай хамтран ажиллах шаардлага тулгарна.

Бүрдүүлсэн их өгөгдөлд машин сургалт, эх хэлний боловсруулалт, гүнзгий сургалтын арга техникүүдийг ашиглан шинжилгээ хийх,

- Машин сургалтын арга техникийг эдийн засгийн төлөвийн тооцоо, тухайн

¹⁵ Data Science

¹⁶ Data Engineer

улирлын болон богино хугацааны төсөөлөл, судалгаа шинжилгээний зорилгоор ашигладаг болох,

- Эх хэлний боловсруулалтын аргыг ашиглан сентимент индекс тооцох, мөнгөний бодлогын шийдвэрт үзүүлж буй нийгмийн сүлжээ хэрэглэгчдийн хариу үйлдлийг шинжлэх,
- Цаашид төв банканд мөнгөний бодлогын DSGE загвар хөгжүүлэхэд агентын зан төлөвт гүнзгий сургалтыг ашиглах,

Монголбанк их өгөгдлийг бодлогын шийдвэр гаргахад шаардагдах мэдээлэл олж авах зорилгоор нийтийн эрх ашгийн төлөө ашиглах болно. Ингэхдээ цаг, хүний нөөц, зардалд хэмнэлттэй зохион байгуулах шаардлага үүснэ. Зохион байгуулалт нь мэдээлэлтэй харьцах үүрэг хүлээсэн мэдээллийн санчдын сүлжээ хэлбэртэй байвал энэ зорилгод нийцнэ.

Дээрх бүх асуудлыг багтаасан их өгөгдлийн удирдлага, зохицуулалтын тухай дунд хугацааны стратегийн баримт бичиг боловсруулан хэрэгжүүлэх нь илүү системтэй ажил болж, улмаар эцсийн үр дүн сайн байх талтай юм.

Ашигласан материал

Aastveit, K. A., Fastbø, T. M., Granziero, E., Paulsen, K. S., & Torstensen, K. N. (2020). Nowcasting Norwegian household consumption with debit card transaction data. *Norges Bank Working Paper Series*.

Aguilar, P., Ghirelli, C., Pacce, M., & Urtasun, A. (2021). Can news help measure economic sentiment? An application in COVID-19 times. *Economic Letters, Volume 199*.

Algaba, A., Borms, S., Kris, B., & Verbeken, B. (2021). Daily news sentiment and monthly surveys: A mixed-frequency dynamic factor model for nowcasting consumer confidence. *Working paper research series No 396, National Bank of Belgium*.

Aprigliano, V., Ardizzi, G., & Monteforte, L. (2019). Using Payment System Data to Forecast Economic Activity. *International Journal of Central Banking*.

Aprigliano, V., Emiliozzi, S., Guitoli, G., Luciani, A., Marcucci, J., & Monteforte, L. (2021). The power of text-based indicators in forecasting the Italian economic activity.

Bank of Korea, B. o. (2021). Establishing a centralized big data processing and integration system. *Knowledge Partnership Program Mongolia*.

Baumgartner, M., & Zahner, J. (2021). Whatever it takes to understand a central banker - embedding their words using neural networks. *Joint Discussion paper series in economics*.

BIS, Doerr, S., Gambacorta, L., & Serena, J. M. (2021). Big data and machine learning in central banking. *BIS Working Papers No 930*.

Broeders, D., & Prenio, J. (2018). *Broeders, Dirk and Jermy Prenio. "Innovative technology in financial supervision (suptech): The experience of early users."* (2018). FSI Insights on policy implementation, No 9.

Chapman, J. T., & Desai, A. (2021). Using Payments Data to Nowcast Macroeconomic Variables During the Onset of COVID-19. *Bank of Canada Staff Working Paper 2021-2*.

Cœuré, B. (2017, 11 24). Policy analysis with big data. *Economic and Financial Regulation in the Era of Big Data*. Paris, France: The Banque de France.

Craig, S. (2021). *The Ultimate Guide to Big Data for Businesses*. Retrieved 2022, from www.techtarget.com: https://media.techtarget.com/digitalguide/images/Misc/EA-Marketing/Eguides/The_Ultimate_Guide_to_Big_Data_for_Businesses_Updated.pdf

D'Amuri, F., & Marcucci, J. (2017). The predictive power of Google searches in

forecasting US unemployment. *International Journal of Forecasting*, Volume 33 Issue 4, 801-816.

Ehrmann, M., & Wabitsch, A. (2022). Central bank communication with non-experts - A road to nowhere? *Journal of Monetary Economics* Volume 127.

Gil, M., Perez, J. J., Fuentes, A. J., & Urtasun, A. (2018). Nowcasting private consumption: traditional indicators, uncertainty measures, credit cards and some internet data. *Banco de Espana Working Paper No. 1842*.

Gorodnichenko, Y., Pham, T., & Talavera, O. (2021). The Voice of Monetary Policy. *Discussion paper, Department of Economics, University of Birmingham*.

Huang, C., Simpson, S., Ulybina, D., & Roitman, A. (2019). News-based Sentiment Indicators. *IMF Working Papers Volume 2019 Issue 273*.

Kiley, M. T. (2020). Financial Conditions and Economic Activity: Insights from Machine Learning. *Financial and Economics Discussion Series, FED*.

Matsumura, K., Oh, Y., Sugo, T., & Takahashi, K. (2021). Nowcasting Economic Activity with Mobility Data. *Bank of Japan Working Paper Series*.

Namsrai, O.-E. (2020). Data Mining. *Smart city project*.

Rooj, D., & Sengupta, R. (2021). Forecasting Private Consumption with Digital Payment Data: A Mixed Frequency Analysis. *ADB Working Paper Series No.1249*.

Sampi, J., & Jooste, C. (2020). Nowcasting Economic Activity in Times of COVID-19. *World Bank Policy Research Working Paper*.

Schubert, A. (2018). AnaCredit: Banking with Pretty Big Data. *Central Banking Focus Report*.

Tian, Z., Hassan, A. F., & Razak, N. (2018). Big Data and SME financing in China. *Journal of Physics: Conference Series*, 1-7.

Verbaan, R., Wilko, B., & Crujisen, C. v. (2017). Using debit card payments data for nowcasting Dutch household consumption. *DNB Working Paper No. 571*.

Vrontos, S. D., Galakis, J., & Vrontos, I. D. (2021). Modelling and predicting U.S. recessions using machine learning techniques. *International Journal of Forecasting*, Volume 37 Issue 2.

Woloszko, N. (2020). Tracking activity in real time with Google Trends. *OECD Economics Department Working papers No.1634*.

МОНГОЛБАНК. (2021). *Төрөөс Мөнгөний Бодлогын Талаар 2022 онд Баримтлах Үндсэн Чиглэл*. Улаанбаатар: Монголбанк.

Цахим эх сурвалжуудын жагсаалт

1. <https://www.oracle.com/big-data/what-is-big-data/>
2. <https://techjury.net/blog/big-data-statistics/#gref>
3. <http://centropolis.mn/machine-learning-юу-хийж-болох-вэ/>
4. <https://www.cloudflare.com/learning/cloud/what-is-the-cloud/>
5. <https://www.investopedia.com/terms/m/mooreslaw.asp>
6. <https://www.lightsondata.com/the-history-of-big-data/>
7. <https://www.lightsondata.com/why-data-governance-is-a-must-for-any-organization/>
8. <https://royalsociety.org/topics-policy/projects/data-governance/>
9. <https://www.techtarget.com/searchdatamanagement/definition/5-Vs-of-big-data>
10. <https://legalinfo.mn/mn>
11. <https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/>
12. <https://bit.ly/3zI3srk>
13. <https://bit.ly/3xl7l4b>

Gil нар. 2018

Эдийн засгийн уламжлалт Испанийн сарын давтамжтай хугацааны цуваан Төлбөрийн картын бодит цагийн тоон индикаторууд дээр төлбөр тооцооны болон картын өгөгдлүүд дээр төлбөр тооцооны болон картын өгөгдөл нь хувийн хэрэглээг сайн тооцооны бодит цагийн гүйлгээний бодит цагийн тоон мэдээлэл, тайлбарлагч бөгөөд стандарт сарын мэдээлэл, интернэт хайлтын тодорхой бус байдлын индексүүд, Google Trend өгөгдлүүдтэй нийсэн үед таамгалал илүү өгөгдөл, тодорхой бус байдлын хайлтын мэдээллийг нэмж, цаг хугацааны сайжирч байжээ. Мөн хайлтын системийн хэмжүүрийг нэмж Испани улсын холимог давтамжтай загвар үнэлэн, Испани мэдээлэл болон тодорхой бус байдлын өрхийн хэрэглээг таамаглах улсын хувийн хэрэглээний тухайн улирлын индексууд нь уламжлалт индикаторуудтай төсөөллийг хийжээ.

Түлхүүр үгс: *private consumption, nowcasting, forecasting, uncertainty, Google Trends*

Aastveit нар. 2020

Дебет картын өндөр давтамжтай Норвег улсын дотоодын бүх дебет картын Дебет картын өгөгдлөөр MIDAS загварыг гүйлгээний өгөгдөл ашиглан гүйлгээний долоо хоногийн давтамжтай тоон өргөтгөснөөр цэгэн болон интервал Норвег Улсын өрхийн хэрэглээг өгөгдлөөр MIDAS загварыг өргөтгөж, өрхийн төсөөлөл нь стандарт бенчмарк загвар болон өндөр давтамжтай хувьсагч бүхий хэрэглээний тухайн улирлын таамаглал бусад загваруудын төсөөллөөс сайжирчээ.

Түлхүүр үгс: *nowcasting, дебет картын картын гүйлгээний өгөгдөл, төсөөллийн үнэмлээ, COVID-19*

Chapman ба Desai.
2021

Төлбөр тооцооны өгөгдөл Канадын төлбөр тооцооны бага болон их дүнтэй Төлбөр тооцооны өгөгдөл болон машин болон машин сургалтын техник бодит цагийн өгөгдөлд машин сургалтын арга сургалтын аргачлал ашигласнаар тухайн ашиглан Канад Улсын макро техникүүд (elastic net, neural network, random улирлын таамаглалын алдаа мэдэгдэхүйц эдийн засгийн таамаглал хийх forest, gradient boosting) ашиглан шинжилж, буурчээ. Төлбөр тооцооны өгөгдлийг цар тахлын үеийн эдийн засгийн төлөвийг машин сургалтаар загварчлан ашигласнаар тодорхойлох, ДНБ, бөөний болон жижиглэн худалдааны ДНБ, бөөний болон жижиглэн худалдааны тус тусын төсөөллийн алдаа бенчмарк худалдааны тухайн улирлын таамаглал хийхэд загвараас 35-40 хувиар, төлбөр тооцооны өгөгдлөөр өргөтгөсөн фактор загвараас 15-20 хувиар буурсан үр дүн гарчээ. Мөн нийт төсөөллийн хугацаанд загварын үр дүн сайжраад зогсохгүй эдийн засгийн хямралын үеүдэд таамаглах чадвар бүр нэмэгддэг хэмээн дүгнэсэн байна.

Arigliano, Ardizzi ба Monteforte. 2019	Төлбөр тооцооны өгөгдөл Итали улсын эдийн засгийн идэвхжиллийг хэмжих давамжтай хувьсагчид бүхий загварт нэмж оруулан, тухайн улирлын ДНБ, түүний дэд бүлгүүдийг тооцоход ашиглажээ.	Итали улсын бага дүнтэй гүйлгээний системийн өгөгдлийг цаг хугацааны холимог давтамжтай хувьсагчид бүхий загварт нэмж оруулан, тухайн улирлын ДНБ, түүний дэд бүлгүүдийг тооцоход ашиглажээ.	Судалгааны үр дүнгээс харахад төлбөр тооцооны өгөгдөл Итали улсын эдийн засгийн үйл хөдлөлийг сайн илэрхийлж байна. Төлбөр тооцооны өгөгдлийг загварт нэмж оруулснаар нийт ДНБ, хэрэглээ, хөрөнгө оруулалт, үйлчилгээний салбарын хувьд загварын тухайн улирлыг төсөөлөх чадвар нэмэгдсэн хэмээн дүгнэжээ.
Rooj болон Sengupta. 2021	Дижитал төлбөр тооцооны өгөгдөл ашиглан Энэтхэг улсын хувийн хэрэглээг таамаглах талаар	Дебет, кредит карт, АТМ гүйлгээ, мобайл банкны гүйлгээний бодит цагийн өгөгдлөөр MIDAS загвар ашиглах нь MIDAS загварыг өргөтгөж, Энэтхэг улсын хувийн хэрэглээг таамаглах чадварыг хувийн хэрэглээний тухайн улирлын таамаглал нэмэгдүүлнэ гэж дүгнэжээ.	Дижитал төлбөр тооцооны өгөгдлөөр өргөтгөсөн MIDAS загвар ашиглах нь хувийн хэрэглээг таамаглах чадварыг нэмэгдүүлнэ гэж дүгнэжээ.
Verbaan, Bolt ба Stuijssen. 2017	Дебет картын өндөр давтамжтай өгөгдөл ашиглан Голланд Улсын өрхийн хэрэглээг таамаглах талаар	Голландын дебет картын гүйлгээний өндөр давтамжтай тоон мэдээллийг MIDAS, DELFI зэрэг загваруудад оруулан үнэлж, төсөөллийн үр дүнг нэгтгэн, Голландын хувийн хэрэглээний тухайн болон дараагийн нэг улирлын төсөөллийг боловсруулжээ.	Дебет картын өгөгдөл нь Голландын өрхийн хэрэглээг загварчлахад үр нөлөөтэй. Дебет картын өгөгдлөөр өргөтгөсөн хамгийн сайн нэгдсэн загварын RMSPE ¹⁷ нь Голландын төв банкны макро эдийн засгийн бодлогын загвар болох DELFI-ын RMSPE-с 18 хувиар бага байна.
D'Amico болон Marcucci. 2017	Гүүгл хайлтын мэдээлэл ашиглан АНУ-ын ажилгүйдлийн таамаглал боловсруулах талаар	Google Trend ашиглан тухайн өдрийн нийт гүүгл хайлтад эзлэх "jobs" хайлтын харьцаагаар Гүүгл индекс (GI) тооцож, энгийн AR(p) загвар ажилгүйдлийн таамаглал боловсруулжээ.	Ажилгүйдлийн сарын төсөөлөл боловсруулах чадварын хувьд GI нь бусад гүүрүүлэгч индикаторуудаас давуу талтай гэж үзсэн. Мөн төсөөлөл боловсруулах чадвар судалгааны үр дүнтэй бусад загвар, түүвэр судалгааны үр дүнтэй харьцуулахад эргэлтийн цэгийг таамаглах чадвар өндөртэй байна.

¹⁷ RMSPE – root mean square prediction error.

- Matsumura, Oh, Sugo болон Takahashi. 2021 Хөдөлгөөнт байдлын үзүүлэлт Гар утас хэрэглэгчдийн 2017.1-2020.03 Хүн амын хөдөлгөөнт байдлын индекс нь ашиглан Япон улсын эдийн хоорондох GPS өгөгдөл ашиглан жижиглэн үйлчилгээний салбарын хэрэглээг бодит засгийн идэвхжиллийг таамаглах худалдаа, үзвэр үйлчилгээ, нийтийн хоол, цагийн торимоор төсөөлөх боломжийг талаар боловсруулах салбарын үйлдвэрлэлийн индекс олгодог байна.
- газруудын хөл хөдөлгөөний индекс Машин тоног төхөөрөмж, сэлбэг хэрэгсэл боловсруулсан. үйлдвэрлэлийн салбарын хөдөлгөөнт байдлын индекс нь тухайн салбарын тухайн улирлын төсөөлөл хийх чадвар сайн гэсэн дүгнэлтэд хүрчээ.
- Sampri болон Jooste. 2020 Хөдөлгөөнт байдлын индекс 2020 оны 2-р сарын 15-ны өдрөөс эхлэлтэй, Хөдөлгөөнт байдлын индекс бүхий MIDAS ашиглан Латин Америк, өдрийн давтамж бүхий “Google mobility index” загвар нь энгийн AR загвараас богино Қарибын тэнгисийн орнуудын үзүүлэлтүүдийг “агаарын бохирдол”-ын түүхэн хугацааны төсөөллийн чадвар өндөр гэж аж үйлдвэрийн салбарын мэдээлэл ашиглан урагш 2019 оны эхэн хүртэл сунгасан. Хөдөлгөөнт байдлын үзүүлэлтүүдээс “Kalman filter” ашиглан ялгасан ерөнхий индекс (фактор)-ийг MIDAS загварт ашиглаж аж үйлдвэрийн салбарын богино хугацааны төлөвийг таамагласан.

Хүснэгт 2. Төв банкны олон нийтийн харилцааны чиглэлээр хийгдсэн судалгааны ажлуудын жишээ

Судлаач	Сэдэв/Улс	Тоон өгөгдөл/Аргачлал	Үр дүн
Ehmann болон Wabitsch. 2021	Европын төв банкны харилцааны үр ашигтай байдлыг ашиглан шинжлэх талаар	Ruhoon программ Банктэй (ЕТБ) холбоотой 5.5 сая гаруй жиргээ агуулгаас хамгаарч иргэдийн хариу үйлдэлд 2 ялгаатай онцлог ажиглагджээ. Үүнд: Твиттер хэрэглэгчийн төв банктай холбоотой 1/ Техникийн шинжтэй мэдээлэлд шинжээчид иргэний тоо, Төв банкны мэдэгдэлд илүүтэй хариу үзүүлэх бөгөөд, Твиттер орчинд тус хариу үйлдэл үзүүлэх хурд зэргээс хамгаарч мэдээллийн талаарх хэлэлцүүлэг богино хугацаанд “шинжээч”, “шинжээч бус (иргэд)” гэх ангилал үргэлжилдэг;	Төв банкаас олон нийгэд түгээж буй мэдээллийн агуулгаас хамгаарч иргэдийн хариу үйлдэлд 2 ялгаатай онцлог ажиглагджээ. Үүнд: Твиттер хэрэглэгчийн төв банктай холбоотой 1/ Техникийн шинжтэй мэдээлэлд шинжээчид иргэний тоо, Төв банкны мэдэгдэлд илүүтэй хариу үзүүлэх бөгөөд, Твиттер орчинд тус хариу үйлдэл үзүүлэх хурд зэргээс хамгаарч мэдээллийн талаарх хэлэлцүүлэг богино хугацаанд “шинжээч”, “шинжээч бус (иргэд)” гэх ангилал үргэлжилдэг;
Gorodnichenko, Pham болон Talavero. 2021	Төв банкуудын бодлогын мэдэгдлээс дууны өнгө, хариултын үеэрх 692 хариултад “neutral network” загварчлал ашиглан Төв банкны нүүрний хувиралд шинжилгээ хийж, төв банкны олон нийттэй харилцах харилцааг шинжих талаар	ЕТБ-ны олон нийгэд чиглэсэн мэдээлэлд Твиттер хэрэглэгч шинжээчид болон иргэд хэрхэн хариу үйлдэл үзүүлж буйг энгийн OLS загвар ашиглан хэмжихийг зорьжээ.	Хэвлэлийн хурлын үеэр “эерэг” дууны өнгөөр олон нийттэй харилцахад: 1/ хөрөнгийн зах зээл дэх хувьцааны үнийг нэмэгдүүлэх нөлөөтэй 2/ Ам.доллар еврогийн эсрэг чангарах нөлөөтэй гэж дүгнэжээ.
Vaughtartner болон Zahner. 2021	Төв банкны бодлогын мэдэгдлийн төрөл бариимт бичигт “Neutral network” ашиглан шинжилгээ хийж, тоон хэлбэрт хувиргах, олон нийттэй харилцах бодлогын үр нөлөөг шинжлэх талаар	Судалгаанд 1980-2021 хооронд BIS-д хэвлэгдсэн 23000 “K-Nearest-Neighbor” аргачлал бүхий 4 төрлийн загвар ашиглан тооцоолол хийсэн.	1/ Хямралын үед ЕТБ-ны олон нийтийн харилцаа нь санхүүгийн зах зээлд оролцогчдын үйл хөдлөлийг намжаахад эерэг нөлөөтэй байна. 2/ Бусад төв банкуудтай харьцуулахад инфляцыг олон нийтийн харилцаа нь RBNZ, FED, болон ECB-тэй төстэй байна.

Хүснэгт 3. Санхүүгийн тогтвортой байдлын шинжилгээнд ашигласан судалгааны ажлуудын жишээ

Судлаач	Сэдэв/Улс	Тоон өгөгдөл/Аргачлал	Үр дүн
Huang, Ulybina, Roitman. 2019	Эдийн засгийн мэдээний болон сентимент үзүүлэлтүүд	“Financial Times” сэтгүүлийн 1980-2019 оны 3 сая гаруй мэдээнд <i>Vector Space Semantics</i> аргачлал ашиглан 20 улс тус бүрд 10 сентимент үзүүлэлт тодорхойлсон.	Судалгааны үр дүнгээс харахад сентимент үзүүлэлтүүд санхүүгийн хямралын өмнө (Бразил 1999), болон эдийн засгийн хүндрэлийн үеэр (Турк 2018) дээд утгатай хүрч байжээ. Мөн сентимент үзүүлэлтүүдийн санхүүгийн хямралыг дохиолох чадвар нь улс орнуудын хувьд харилцан адилгүй гэж дүгнэсэн. Тухайлбал, Бразилулсад “fear” болон “hedging”, Мексик болон Швед улсад “crisis”, Аргентин болон Боливи улсад “opinion” сентиментүүд сайн дохиологч болж байжээ.
D. Vrontos нар. 2020	Машин сургалтын зах зээлийн сарын техник ашиглан АНУ эдийн засгийн хямралыг таамаглах нь	Макро эдийн засаг, санхүүгийн зах зээлийн сарын давтамжтай 56 хувьсагч болон машин сургалтын техникүүд (CART recursive partitioning algorithm, bagging, random forest, adaptive boosting, gradient boosting, naïve bayes, k-nearest neighbors) ашиглан төсөөлөл боловсруулах загварчлал хөгжүүлсэн.	Машин сургалтын техник бүхий шинэ загварууд энгийн Logit/Probit загваруудаас богино, дунд, урт хугацаанд таамаглалын чадвар өндөргүй байжээ.
Kiley. 2020	Санхүүгийн нөхцөл, эдийн засгийн идэвхжилийг хэмжихэд машин сургалтыг ашиглах нь	Санхүү, эдийн засгийн сарын давтамж бүхий хувьсагчид болон машин сургалтын (LASSO, elastic net) аргачлалууд ашиглан санхүүгийн нөхцөлийн индекс (ML-FCI) тоолсон. Үнэлгээний аргачлал нь энгийн регрессгүй төстэй боловч үнэлгээнд ашиглагдаж буй хувьсагчдын тоо олон байсан. Тиймээс үнэлгээний хэмжээсийг бууруулах зорилгоор машин сургалтын техникүүд ашиглажээ.	Судалгааны үр дүнгээс харахад ML-FCI индекс нь 1990 болон 2000-аад оны эхний үеийн хямралуудаас урьдаж санхүүгийн нөхцөл муудахыг дохиолож байсан нь Чикагогийн Холбооны Нөөцийн Банкны “National Financial Conditions Index”-ээс давуу байна гэж үзжээ.

Broeders, нар. 2018	Хяналтын (surtech)-ийн түүнийг хяналт, ашиглах судлах нь	Хяналтын технологи тухай, санхүүгийн шалгалт ирээдүйд зорилготой. Тухайлбал, дэх задгай их өгөгдөлд замаар макро болон хяналт, шалгалтыг	Хяналтын технологи (surtech) нь сүүлийн үеийн дотоод мэдээлэлд суурилсан арилжаа, луйврын гүйлгээг илрүүлэх, буй байгууллагын хөрвөх чадварын эрсдэлийг цаг алдалгүй тодорхойлох, үл хөдлөх хөрөнгийн зах зээлийн нөхцөл өгөгдлийг нэгтгэх, баталгаажуулах, шинжилгээ хийх өгөгдлийг нэгтгэх, баталгаажуулах, шинжилгээ хийх замаар макро болон микро захистой бодлогын хүрээнд хяналт, шалгалтыг дэмждэг байна.	Surtech-ийн нийтлэг хэрэглээ нь дотоод мэдээлэлд суурилсан арилжаа, луйврын гүйлгээг илрүүлэх, буй байгууллагын хөрвөх чадварын эрсдэлийг цаг алдалгүй тодорхойлох, үл хөдлөх хөрөнгийн зах зээлийн нөхцөл байдлыг таамаглах зэрэг байна. Surtech-г ашигласнаар зардал буурч, үр өгөөж нэмэгдэх, хүчин чадал тэлэх давуу талтай.
Schubert. 2018	Зээлийн мэдээллийн сан, Санхүүгийн шалгалтад ашиглах судлах нь	Европын төв банк бусад Төв банкуудтай хамтран “дэлхийн зээлийн бүртгэл” бүхий “Анакредит” нэгдсэн санг боловсруулж байна. Энэ нь микро түвшинд зээл, зээлийн эрсдэлийн их өгөгдлийг бүрдүүлэн, хадгалах түүнд дүн шинжилгээ хийж, Төв банк болон банкууд зээлийн мэдээллийн болон хяналт, шалгалтад ашиглах боломж бүхий нэгдсэн систем юм.	Surtech нь өгөгдлийг цуглуулах хүрээнд API, үүлэн тооцоолол, чагбог, data input approach, data pull approach, машин-уншихуйц зохицуулалт (machine-readable regulation) зэргийг ашигладаг. Харин Өгөгдлийн шинжилгээний хүрээнд хиймэл оюун ухаан, машин сургалт, эх хэлний боловсруулалт, дүрс танилт, ньюрел нетворк, рандом форест зэрэг техникүүдийг ашиглаж байна.	Анакредит нь зээлийн болон зээлийн өсөлтөөс үүсэх эрсдэлийг илэрхийлэгч үзүүлэлтүүдээс бүрддэг бөгөөд энэ нь бодлого боловсруулагчид зээлийн өсөлтийг хянах, банкны зээлийн өсөлтийг хязгаарлах чиглэлийн макро захистой бодлогын шийдвэрүүдийн үр дүнг үнэлэхэд дэмжлэг үзүүлнэ.