

# ДААТГАЛЫН БАЙГУУЛЛАГЫН ИХ ӨГӨГДӨЛД ӨГӨГДЛИЙН УУРХАЙН АРГУУД АШИГЛАН ШИНЭ МЭДЛЭГ НЭЭХ: СУДАЛГАА, ХЭРЭГЖҮҮЛЭЛТ

Г.Баттөгөлдөр#1, Д.Намчин#2, Н.Эрдэнэтуяа#3, Н.Оюун-Эрдэнэ#4,  
#1, #2, #4 Мэдээлэл компьютерийн ухааны тэнхим, Хэрэглээний Шинжлэх ухаан Инженерчлэлийн  
Сургууль, Монгол Улсын Их Сургууль  
#3 Бизнесийн Удирдлага Хүмүүнлэгийн Сургууль, Шинжлэх Ухаан Технологийн Их Сургууль  
#1 tuugii83@gmail.com, #2 namchin@gmail.com, #3 erdenetuya2080@gmail.com,  
#4 oyunerdene@seas.num.edu.mn

Хураангуй — HeavenPro програм хангамж нь даатгалын байгууллагад ашиглагддаг үндсэн систем бөгөөд бид уг системийг 2010 оноос хойш хөгжүүлэн одоогоор даатгалын байгууллагууд өргөнөөр ашиглаж байна. HeavenPro систем нь даатгалын байгууллагатай холбоотой өгөгдлийг үүсгэн хадгалж өөрийн өгөгдлийн санг бий болгодог. Энэ хуримтлагдсан өгөгдлөөс бизнесийн байгууллагын өнөөгийн болон хэтийн төлөвт хэрэг болохуйц хэрэгтэй мэдлэг мэдээлэл гаргаж авахад өгөгдлийн уурхай буюу өгөгдлийн тандалтын аргуудыг ашиглан шинэ мэдлэг мэдээлэл гаргаж авах хэрэгцээ шаардлага бий болсон. Судалгааны ажлын гол зорилго нь даатгалын байгууллагын системийн Албан журмын жолоочийн хариуцлагын даатгал дээрх нөхөн төлбөрийн өгөгдөлд өгөгдлийн тандалтын холбоо хамаарлыг илрүүлэх Argiori арга ашиглан дүн шинжилгээ хийж далд байгаа өгөгдлийн хэв маягуудыг тодруулах, параметр хоорондын харилцан хамаарлыг нээн илрүүлэхэд чиглэгдсэн. Өгөгдлийн холбоо хамаарлыг тодорхойлох үндсэн арга болох Argiori-ийг HeavenPro даатгалын үндсэн системд хөгжүүлэн нэвтрүүлж даатгалын нөхөн төлбөрийн тээврийн хэрэгслийн өгөгдөлд дүн шинжилгээ хийж бизнесийн байгууллага болон хэрэглэгчдэд зориулсан шинэ мэдлэг, мэдээлэл гаргасан нь даатгалын байгууллага, хэрэглэгч харилцан ашигтай ажиллахад тодорхой хувь нэмэр оруулна гэдэгт итгэж байна.

*Түлхүүр үг: Өгөгдлийн тандалтын Apriori арга, нөхөн төлбөрийн өгөгдөл, даатгалын өгөгдлийн сан, өгөгдлийн холбоо хамаарал, HeavenPro даатгалын үндсэн систем*

## I. УДИРТГАЛ

HeavenPro даатгалын үндсэн системийг Монголын томоохон даатгалын байгууллагууд ашиглаж байгаа ба даатгалын байгууллагын системд даатгалын их өгөгдөл цугларан хадгалагдаж байна. HeavenPro систем нь дараах модулиудаас бүрддэг. Үүнд: Серверийн аппликейшн, Мэдээллийн бааз, Харилцагч, Параметр, Дүрэм, Policy тохиргоо, Санхүү, Ерөнхий дэвтэр, Үндсэн хөрөнгө, Бараа материал, Админ удирдлага, Балансын гадуурх, Даатгалын гэрээ, Даатгалын хэлцэл болон гэрээт баталгаа, Няравын бүртгэл, Дуудлага, Нөхөн төлбөр, Тайлан, Лог бүртгэл, Өдөр өндөрлөлт, Давхар даатгалын гэрээ, Давхар даатгалын хэлцэл, Давхар

даатгалын нөхөн төлбөр, Онлайн даатгал гэх мэт 40 хүртэлх модультай.

HeavenPro даатгалын үндсэн системд хуримтлагдсан өгөгдөлд тулгуурлан судалгааны эхний шатанд Албан журмын жолоочийн хариуцлагын даатгал дээрх нөхөн төлбөр авсан харилцагчийн нас, хүйс, тохиолдол гарсан аймаг/хот, сум/дүүргийн байршил, огноо болон тээврийн хэрэгсэлтэй холбоотой нийт 5206 бичлэг, 12 талбар бүхий өгөгдлийн багц дээр анализ хийсэн.

Монголд даатгалын байгууллагын системийн их өгөгдөлд өгөгдлийн уурхайн алгоритм болон арга техник ашиглан тандалт хийсэн судалгааны ажил одоогийн байдлаар хараахан хийгдээгүй байна.

## II. ӨГӨГДЛИЙН УУРХАЙ БУЮУ ТАНДАЛТ

Их өгөгдөлд задлан шинжилгээ хийж, шийдвэр гаргахад туслах мэдээллийг олж илрүүлэх үйл явцыг Өгөгдлийн уурхай буюу тандалт гэнэ. Өөрөөр хэлбэл аливаа байгууллага үйл ажиллагаатайгаа холбоотой мэдээллүүдийг хадгалсаар, их өгөгдлийн архивыг үүсгэсэн байдаг бөгөөд түүнийг зөвхөн тайлан тооцоо гаргахад ашигладаг байна. Гэтэл их өгөгдлийн цаана маш олон сонирхолтой, цаашдын үйл ажиллагаандаа ашиглаж болохуйц мэдлэг, мэдээлэл нуугдаж байдаг. Үүнийг SQL асуулгаар гаргаж ирэх боломжгүй бөгөөд математикийн тусгай тооцооллоор илрүүлдэг байна. Үүнээс үүдэн SQL асуулгаар гаргаж авах боломжгүй мэдээллийг илрүүлэх математикийн тусгай аргууд бий болсон ба үүнийг өгөгдлийн тандалтын техникүүд гэдэг. Эдгээр техникүүд нь өгөгдлийг өөр өөр өнцгөөс харж тооцоолол хийж шаардлагатай үр дүнг гаргадаг. Өгөгдлийн тандалтын арга техникийг хэрэгжүүлэх эхний алхам нь сонгосон өгөгдөлд урьдчилсан боловсруулалт хийх арга буюу Data Preprocessing юм.

Өгөгдөлд урьдчилсан боловсруулалт хийх үндэслэл нь:

- Сонгосон өгөгдлүүд ихэнхдээ дараах өгөгдлүүдийг агуулдаг :
- *Бүрэн бус, дутуу өгөгдөл* : шинж чанар болон утгын хувьд бүрэн бус, дутуу өгөгдөл

- Алдаатай оруулсан өгөгдөл: хязгаараас хэт давсан утга агуулсан.
- Нийцгүй өгөгдөл : Код дугаарлалт эсвэл нэр нь зөрүүтэй утга агуулсан өгөгдлүүд байдаг ба өгөгдөлд урьдчилсан боловсруулалт хийх дараах аргуудтай:
- Өгөгдлийг цэвэрлэх {Data cleaning}
- Өгөгдлийг нэгтгэх {Data Integration}
- Өгөгдлийн хувиргалт {Data Transformation}
- Өгөгдлийн хэмжээг багасгах {Data Reduction}



Зураг №1 Өгөгдлийг урьдчилан боловсруулах үе шат (Data Preprocessing)

Өгөгдлийг урьдчилан боловсруулах алхам нь өгөгдлийн тандалтын процессийн эхний чухал алхам бөгөөд эндээс бэлэн болсон өгөгдөлд өгөгдөл тандалтын аргуудыг ашиглан үр дүн гаргана.

Өгөгдлийн уурхайн дараах үндсэн аргууд байдаг.

**A. Өгөгдлийн тандалтын гаж өгөгдлийг илрүүлэх арга {Anomaly detection}:**

Их өгөгдөлд дүн шинжилгээ хийхэд үр дүн буруу гарахад нөлөөлөх бичлэгийг олж илрүүлэх арга. Жишээ нь: Даатгалын нөхөн төлбөрийн өгөгдлүүдээс санаатайгаар нөхөн төлбөр авч буй өгөгдлийг илрүүлэх боломжтой.

**B. Өгөгдлийн тандалтын ангилах арга {Classification}:**

Шинээр орж ирэх өгөгдлийг өмнө үүсгэн сургасан сургалтын багц дээр шалгаж ангилал хийх арга юм. Жишээ нь : Даатгалын байгууллагын харилцагчийн мэдээлэлд ангилал хийснээр худалдан авалт хийж буй шинэ харилцагчийг өмнөх харилцагчийн ангиллын дагуу ангилах боломжтой болно.

**C. Өгөгдлийн тандалтын бүлэглэх арга {Clustering} :**

Өгөгдөл болон хувсагчдын ижил төстэй байдлаар бүлэглэж хуваах арга юм. Уг арга нь статистикийн дүн шинжилгээ хийх, мэдээлэл сэргээх, хэв маяг илрүүлэх, дүрс шинжилгээ хийх, өгөгдөл шахах болон компьютер графикт ашиглагдаж байна.

Жишээ нь: Даатгалын байгууллагын харилцагчийн нас, хураамжийн орлого болон нөхөн төлбөрийн дүнг бүлэглэн тодорхой үр дүн гаргаж болно.

**D. Өгөгдлийн тандалтын холбоо хамаарлыг илрүүлэх арга {Association rule learning}:**

Энэ нь өгөгдөл, хувьсагч хоорондын харилцаа холбоог илрүүлэх арга юм. Холбоо хамаарлын дүрмийг support, confidence гэсэн үндсэн 2 хэмжигдэхүүнээр хэмждэг.

- Support - Зүйлүүдийн олонлог нь нийт өгөгдлийн олонлогийн хэдэн хувьд нь илэрч байгааг олно. Жишээ нь: {Тээврийн хэрэгслийн даатгал, Гэнэтийн ослын даатгал}
- Confidence -  $\{X\} \Rightarrow \{Y\}$  {X Зүйл болон зүйлүүдийг агуулсан олонлогт Y зүйл хэдэн хувьд нь орсныг илэрхийлэх хувь}
- Жишээ нь: {Тээврийн хэрэгслийн даатгал, Гэнэтийн ослын даатгал}  $\Rightarrow$  {Албан журмын даатгал}

Даатгалын байгууллагын харилцагчдын даатгалын бүтээгдэхүүн худалдан авалтын холбоо хамаарлыг тодорхойлсноор байгууллагын борлуулалт болон нөхөн төлбөрт нөлөөлж болохуйц хэрэгтэй мэдлэг мэдээллийг гаргах боломжтой болно. Холбоо хамаарлыг илрүүлэх арга нь их өгөгдлүүдийн хоорондын холбоо хамаарлыг нээн илрүүлэхэд ашиглагддаг. Уг аргыг бизнесийн салбарт ихээхэн ашигладаг бөгөөд өгөгдлүүдийн холбоо хамаарлыг нээн илрүүлснээр бизнесийн байгууллагын өнөөгийн болон хэтийн төлөвт хэрэг болохуйц байгууллагын шийдвэр гаргах түвшний албан тушаалтнуудад зориулсан ашигтай хэрэгтэй мэдээлэл гаргаж өгдөг.

Холбоо хамаарлын дүрмийг тодорхойлох үндсэн аргуудаас Apriori алгоритмыг сонгон авч судалгааны ажилдаа ашигласан.

**E. Apriori алгоритм**

Apriori алгоритм нь холбоо хамаарлын дүрмийг судлахад ашиглагддаг сонгодог алгоритм юм.

Тус алгоритм нь өгөгдлийн сангийн зүйлүүдийн олонлогийг тоолон боломжит дүрмүүдийг хайж цааш нэр дэвшүүлэх байдлаар ажилладаг. Өөрөөр хэлбэр хамгийн бага support-ийг хангасан зүйлүүдийн олонлогийг түүвэрлэн авч confidence-ийг тооцох зарчмаар ажилладаг. Хэрэв тус confidence нь хамгийн бага confidence-ийг хангасан бол дүрмийг нэр дэвшүүлдэг ба confidence өндөр байх тусам дүрэм хүчтэй байдаг. Apriori алгоритмын pseudo код буюу хуурмаг хэлийг Зураг 2-т үзүүлэв.

```

Apriori(T, ε)
L1 ← {large 1 - itemsets}
k ← 2
while Lk-1 ≠ ∅
    Ck ← {a ∪ {b} | a ∈ Lk-1 ∧ b ∉ a} - {c | {s | s ⊆ c ∧ |s| = k - 1} ⊄ Lk-1}
    for transactions t ∈ T
        Ct ← {c | c ∈ Ck ∧ c ⊆ t}
        for candidates c ∈ Ct
            count[c] ← count[c] + 1
    Lk ← {c | c ∈ Ck ∧ count[c] ≥ ε}
    k ← k + 1
return ∪k Lk
    
```

Зураг №2 Apriori алгоритмын pseudo код буюу хуурмаг хэлийг үзүүлэв

Apriori алгоритмыг тайлбарласан даатгалын байгууллагын өгөгдлийн жишээг хүснэгт 1-д харуулав.

Хүснэгт №1 Даатгалын байгууллагын борлуулалтын өгөгдөл

№	Албан журам	Тээврийн хэрэгсэл	Гэнэтийн осол	Эрүүл мэнд	Хариуцлага
1	1	1			
2			1	1	
3		1	1		1
4	1	1	1		
5	1	1	1		1

Хүснэглт №1-ийн мэдээллээс дараах үр дүнг гаргаж болно. Үүнд:

Албан журмын даатгалыг худалдан авсан харилцагч тээврийн хэрэгслийн даатгалыг худалдан авдаг.

Дээрх өгөгдлийг HeavenPro даатгалын систем дээр хөгжүүлсэн Apriori арга, WEKA програм хангамжийн Apriori арга дээр ажиллуулж туршсан үр дүнгийн харьцуулалтыг хүснэгт 2-т үзүүлбэ.

Хүснэгт №2. WEKA APRIORI БОЛОН HEAVENPRO APRIORI АРГЫН ХАРЬЦУУЛАЛТ

№	Rules	WEKA.Apriori		Our.Apriori	
		Support	Confidence	Support	Confidence
1	{Хариуцлага} => {Гэнэтийн осол}	40%	100%	40%	100%
2	{Хариуцлага} => {Тээврийн хэрэгсэл}	40%	100%	40%	100%
3	{Гэнэтийн осол, Албан журам} => {Тээврийн хэрэгсэл}	40%	100%	40%	100%
4	{Хариуцлага} => {Гэнэтийн осол, Тээврийн хэрэгсэл}	40%	100%	40%	100%
5	{Хариуцлага, Тээврийн хэрэгсэл} => {Гэнэтийн осол}	40%	100%	40%	100%
6	{Гэнэтийн осол, Хариуцлага} => {Тээврийн хэрэгсэл}	40%	100%	40%	100%
7	{Албан журам} => {Тээврийн хэрэгсэл}	60%	100%	60%	100%

Хүснэгт №2-оос харахад хоёр програм дээр ажиллуулсан Apriori алгоритмийн үр дүн ижил гарсан.

### III. ХЭРЭГЖҮҮЛЭЛТ

HeavenPro даатгалын байгууллагын үндсэн систем дээр өгөгдлийн тандалтын Apriori алгоритм аргыг хөгжүүлж нэвтрүүлэлт хийсэн. Зураг №3 үзүүлэв.

Уг систем нь Client-Server технологи дээр суурилсан Oracle өгөгдлийн баазтай систем юм.

- Server тал нь windows .net c#, php хэл дээр бичигдсэн
- Client тал: windows .net c#, html, JavaScript, CSS хэл дээр бичигдсэн
- Client JavaScript framework: jQuery – jqwidgets

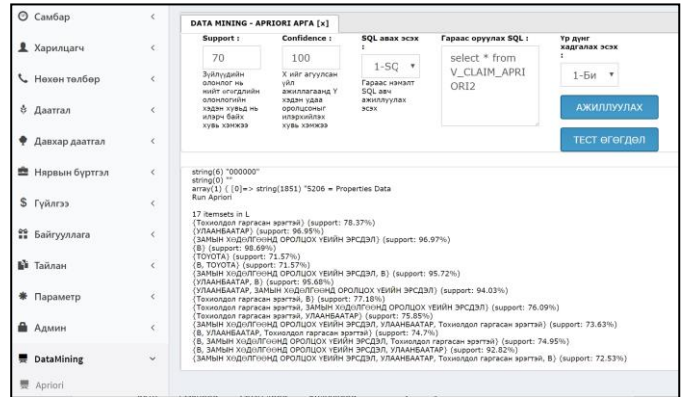
Oracle өгөгдлийн бааз нь холбоост өгөгдлийн сангийн удирдах систем (RDBMS) юм.

Apriori алгоритмыг систем дээрээ сервер талдаа буюу windows .net c# хэл дээр хөгжүүлсэн.

Энэ нь client талаас тооцоололд хэрэгтэй support болон confidence-ийн хувь зэрэг өгөгдлүүдийг хэрэглэгчээс авч сервер програмуу Socket-оор мэдээллийг илгээнэ. Сервер компьютер дээр байрлах сервер програм нь client-ээс хүсэлтийг хүлээн аваад өөр дээрээ өгөгдлийн тандалтын Apriori алгоритмыг ажиллуулж тооцоолол хийх бөгөөд гарсан үр дүнг

Oracle мэдээллийг баазруу хадгалах ба Client хэсэг рүү мэдээллийг давхар илгээнэ.

Client хэсэг дээр гарсан үр дүнг хүлээн авч мэдээллийг харуулна. Мөн мэдээллийн баазад хадгалагдсан үр дүнгээр хэрэглэгчдэд ойлгомжтой байдлаар график дүрслэл, хүснэгт зэргээр визуаль хэлбэрт шилжүүлэн тайлан мэдээ байдлаар гаргах боломжтой юм.



Зураг №3 HeavenPro даатгалын үндсэн системийн дэлгэц

#### A. Өгөгдөлд урьдчилан боловсруулалт хийх нь :

Apriori алгоритмыг ашиглан HeavenPro систем ашигладаг “XXX” даатгалын компанийн Жолоочийн хариуцлагын албан журмын даатгал дээр гарсан нөхөн төлбөрийн 2015–2017 оны нийт 5206 бичлэг, 12 талбар бүхий бодит өгөгдөл дээр дүн шинжилгээ хийсэн. Талбарын авсан утгууд:

- Даатгалын харилцагчийн хүйс
- Даатгалын харилцагчийн нас, бүлэглэснээр
- Тохиолдол гаргасан тээврийн хэрэгслийн үйлдвэр
- Тохиолдол гаргасан тээврийн хэрэгслийн модель
- Тохиолдол гаргасан тээврийн хэрэгслийн ангилал
- Тохиолдол гаргасан жолоочийн хүйс
- Тохиолдол гаргасан жолоочийн нас, бүлэглэснээр
- Тохиолдол гарсан улирал
- Тохиолдол гарсан өдрийн цагийн интервал
- Тохиолдол гарсан аймаг/хотын нэр
- Тохиолдол гарсан сум/дүүргийн нэр
- Тохиолдсон эрсдэлийн нэр

Өгөгдлийн урьдчилан боловсруулалт нь өгөгдлийн тандалтын процессийн чухал алхам боловч ихэнхдээ энэ алхамыг бүрэн гүйцэд ашиглаж чаддаггүй. Өгөгдөлд урьдчилан боловсруулалт хийснээр гарах үр дүн нь илүү бодит байна. Даатгалын байгууллагын өгөгдлөөс анхны гарсан

мэдээлэлд дараах өгөгдлийн урьдчилсан боловсруулалт хийгдсэн.

HeavenPro даатгалын үндсэн системийн өгөгдлийн баазын хүснэгт {table}-үүдээс холбогдох нөхөн төлбөрийн мэдээллийг SQL асуулгаар шүүж гаргасан. Шүүлт хийсэн гол талбарууд:

- Жолоочийн хариуцлагын албан журмын даатгалын бүтээгдэхүүн байх
- Нөхөн төлбөр олгосон байх
- Тохиолдол гаргасан он нь 2015 он болон түүнээс их байх
- Харилцагч нь иргэн байх

Уг шүүлтүүдээс гарч ирсэн өгөгдлүүд дээр дараах урьдчилсан боловсруулалтуудыг хийсэн.

- Харилцагчийн төрсөн огнооны мэдээлэл дутуу байгаа харилцагчийн мэдээллийг Монгол улсын иргэний регистрийн дугаараас төрсөн огноог гаргаж авсан.
- Гадаад улсын иргэний төрсөн он оруулаагүй өгөгдлийг нөхөх боломжтойг нөхөж боломжгүй байсан өгөгдлийг хассан.
- Нөхөн төлбөрийн мэдээлэлд тохиолдол гаргасан жолоочийн мэдээлэл холбогдоогүй бол уг нөхөн төлбөр гарсан даатгалын гэрээний үндсэн харилцагчийн мэдээллийг нөхөж авсан.
- Гарч ирсэн мэдээллийг үндэслэн даатгалын харилцагч болон тохиолдол гаргасан харилцагчийн насыг бүлэглэсэн.
- Тохиолдол гаргасан огноог улирлаар ангилсан. (Хавар, Зун, Намар, Өвөл)
- Тохиолдол гаргасан цагийг ангилсан. (01-06, 07-12, 13-18, 19-24)
- {Missing Value} – {Data cleaning} Хоосон утгыг тооцоололд оруулах эсэх сонголтоор өгөгдлийг шүүсэн.

Уг өгөгдлийг Oracle мэдээллийн базаас авч HeavenPro даатгалын үндсэн систем дээр хөгжүүлсэн. Apriori алгоритмыг 2 удаа ажиллуулж боловсруулалт хийсэн. Хүснэгт №3, №4 - өөр үзүүлэв. Өөрөөр хэлбэл нийт өгөгдөл дээр анхны удаа Apriori алгоритмыг ажиллуулан гарсан дүрмүүдээс хамгийн өндөр support болон confidence-тэй өгөгдлүүдийг дахин шүүж авсан. Уг дахин шүүж авсан өгөгдлөөс өмнөх дүрэмд илэрсэн талбаруудыг хасаад бусад илрээгүй талбарууд дээр дахин Apriori алгоритмыг ажиллуулж өмнөх алхамаас илүү нарийн үр дүнг гаргасан.

Хүснэгт №3 APRIORI АРГЫГ ЭХНИЙ УДААД АЖИЛЛУУЛСАН

CUSTSEX	CUSTAGE	CUSTO...	C...	CUSTO...	ACCAGE	ACCSEX	ACCIDENT...	ACCTIMEI...	CITYNAME	DISTRICTNAME	RISKNAME		
Эрэгтэй	...	A25-35 настай			A25-35	...	Эрэгтэй	Өвөл(11-1)	07-12 цаг	УЛААНБААТАР	БАНЗУРХ ДҮҮРЭГ	ЗАМЫН ХӨДӨЛГӨӨНД	
Эрэгтэй	...	A35-40 настай	HYUNDAI	B	STAREX	A35-40	...	Эрэгтэй	Хавар(2-4)	07-12 цаг	УЛААНБААТАР	СҮХБААТАР ДҮҮРЭГ	ЗАМЫН ХӨДӨЛГӨӨНД
Эрэгтэй	...	A50 аас дээш			A50 аас...	...	Эрэгтэй	Хавар(2-4)	13-18 цаг	ОРХОН	БАН ӨНДӨР ХОРО	ЗАМЫН ХӨДӨЛГӨӨНД	
Эрэгтэй	...	A40-50 настай	HONDA	B		A40-50	...	Эрэгтэй	Хавар(2-4)	07-12 цаг	УЛААНБААТАР	БАНЗУРХ ДҮҮРЭГ	ЗАМЫН ХӨДӨЛГӨӨНД
Эрэгтэй	...	A25-35 настай	TOYOTA	B	PRJUS	A25-35	...	Эрэгтэй	Хавар(2-4)	19-24 цаг	УЛААНБААТАР	ЧИНГЭЛТЭЙ ДҮҮРЭГ	ЗАМЫН ХӨДӨЛГӨӨНД
Эрэгтэй	...	A25-35 настай	TOYOTA	B	PRJUS	A25-35	...	Эрэгтэй	Хавар(2-4)	07-12 цаг	УЛААНБААТАР	ЧИНГЭЛТЭЙ ДҮҮРЭГ	ЗАМЫН ХӨДӨЛГӨӨНД
Энэгтэй	...	A25-35 настай	TOYOTA	B		A25-35	...	Энэгтэй	Хавар(2-4)	07-12 цаг	УЛААНБААТАР	ЧИНГЭЛТЭЙ ДҮҮРЭГ	ЗАМЫН ХӨДӨЛГӨӨНД
Эрэгтэй	...	A40-50 настай	TOYOTA	B		A40-50	...	Эрэгтэй	Хавар(2-4)	07-12 цаг	УЛААНБААТАР	БАНЗУРХ ДҮҮРЭГ	ЗАМЫН ХӨДӨЛГӨӨНД
Эрэгтэй	...	A25-35 настай	TOYOTA	B	PRJUS	A25-35	...	Эрэгтэй	Хавар(2-4)	13-18 цаг	УЛААНБААТАР	БАНЗУРХ ДҮҮРЭГ	ЗАМЫН ХӨДӨЛГӨӨНД
Эрэгтэй	...	A25-35 настай	TOYOTA	B	PRJUS	A25-35	...	Эрэгтэй	Хавар(2-4)	13-18 цаг	УЛААНБААТАР	БАНЗУРХ ДҮҮРЭГ	ЗАМЫН ХӨДӨЛГӨӨНД
Эрэгтэй	...	A25-35 настай	HONDA	B		A25-35	...	Эрэгтэй	Хавар(2-4)	13-18 цаг	УЛААНБААТАР	БАНЗУРХ ДҮҮРЭГ	ЗАМЫН ХӨДӨЛГӨӨНД
Эрэгтэй	...	A25-35 настай				A25-35	...	Эрэгтэй	Зун(5-7)	13-18 цаг	УЛААНБААТАР	СҮХБААТАР ДҮҮРЭГ	ЗАМЫН ХӨДӨЛГӨӨНД
Эрэгтэй	...	A25-35 настай	SUBARU	B		A25-35	...	Эрэгтэй	Хавар(2-4)	13-18 цаг	УЛААНБААТАР	ЧИНГЭЛТЭЙ ДҮҮРЭГ	ЗАМЫН ХӨДӨЛГӨӨНД

Хүснэгт №4 APRIORI АРГЫГ 2 ДАХЬ УДАА АЖИЛЛУУЛСАН ӨГӨГДӨЛ

ACCAGE	CUSTOMERMARK	ACCIDENTDATE	ACCTIMEINTERVAL	DISTRICTNAME
25-35 настай	PRJUS	Хавар(2-4)	19-24 цаг	ЧИНГЭЛТЭЙ ДҮҮРЭГ
25-35 настай	PRJUS	Хавар(2-4)	19-24 цаг	ЧИНГЭЛТЭЙ ДҮҮРЭГ
40-50 настай		Хавар(2-4)	07-12 цаг	БАЯНГОЛ ДҮҮРЭГ
25-35 настай	PRJUS	Хавар(2-4)	13-18 цаг	БАЯНГОЛ ДҮҮРЭГ
25-35 настай	PRJUS	Хавар(2-4)	13-18 цаг	БАЯНГОЛ ДҮҮРЭГ
50 аас дээш настай		Хавар(2-4)	07-12 цаг	ЧИНГЭЛТЭЙ ДҮҮРЭГ
50 аас дээш настай	LAND CRUISER 200	Хавар(2-4)	07-12 цаг	БАЯНГОЛ ДҮҮРЭГ
40-50 настай	PRJUS	Зун(5-7)	07-12 цаг	СҮХБААТАР ДҮҮРЭГ
35-40 настай	CHASER	Хавар(2-4)	19-24 цаг	БАЯНГОЛ ДҮҮРЭГ
40-50 настай	PRJUS	Өвөл(11-1)	13-18 цаг	СОНГИНО ХАЙРХАН ДҮҮРЭГ

ҮР ДҮН

Судалгааны эхний шатанд даатгалын байгууллагуудад ашиглагдаж буй HeavenPro даатгалын үндсэн систем дээр өгөгдлийн уурхайн холбоо хамаарлыг илрүүлэх Apriori аргыг хөгжүүлэн нэвтрүүлсэн. Уг аргыг HeavenPro систем дотроос сонгож орох бөгөөд уг аргыг өгөгдөл болон гарсан үр дүн дүрмүүдээс хамааруулан 2 удаа Apriori аргыг ажиллуулсан үр дүнг гаргасан.

A. Алхам №1: Apriori

Support: 50%

Confidence: 100%

Нийт 5206 бичлэг 12 талбар бүхий даатгалын нөхөн төлбөрийн өгөгдлийн багцад өгөгдлийн урьдчилсан боловсруулах шатны хоосон утгыг хасаж шүүхэд Apriori аргыг эхний удаа ажиллуулж дараах хамгийн өндөр Support болон Confidence - тэй холбоо хамаарлын гол дүрмийг гарган авсан.

- Дүрэм : {B, УЛААНБААТАР, TOYOTA, Эрэгтэй} => {ЗАМЫН ХӨДӨЛГӨӨНД ОРОЛЦОХ ҮЕИЙН ЭРСДЭЛ} (support: 54.55%, confidence: 100%)

Эндээс “УЛААНБААТАР хотод TOYOTA маркын B ангиллын авто тээврийн хэрэгслийг эрэгтэй жолооч замын хөдөлгөөнд жолоодож оролцох үедээ ЗАМЫН ХӨДӨЛГӨӨНД ОРОЛЦОХ ҮЕИЙН ЭРСДЭЛ-ийг хамгийн их гаргадаг” гэсэн холбоо хамаарал гарсан.

B. Алхам №2: Apriori

Эхний алхамаар гарсан дүрмээс хамааруулан дараах шүүлтийг хийсэн. Анхны 5206 бичлэгээс эхний дүрмээр гарсан талбаруудын утгуудаар шүүлт хийгээд эхний дүрмээр илэрсэн талбаруудаас бусад талбаруудыг сонгон авч Apriori аргыг дахин хэрэглэсэн.



Support: 10%

Confidence: 40%

Хоосон утгыг оруулахгүй

- Дүрэм №1: {БАЯНЗҮРХ ДҮҮРЭГ} => {PRIUS} (support: 12.01%, confidence: 51.98%)
- Дүрэм №2: {БАЯНГОЛ ДҮҮРЭГ} => {PRIUS} (support: 12.75%, confidence: 56.81%)
- Дүрэм №3: {Өвөл(11-1), 25-35 настай} => {PRIUS} (support: 10.35%, confidence: 56.83%)
- Дүрэм №4: {19-24 цаг} => {PRIUS} (support: 10.79%, confidence: 53.7%, lift: 1.04%)
- Дүрэм №5: {07-12 цаг} => {PRIUS} (support: 18.69%, confidence: 51.26%, lift: 0.99%)

Эндээс “Баянзүрх дүүрэгт 25-35 насны жолооч нар 19-24 цагийн хооронд их тохиолдол гаргадаг” гэсэн дүрмийг гарган авсан.

#### ДҮГНЭЛТ

Одоогийн байдлаар зарим даатгалын байгууллагууд нь HeavenPro даатгалын үндсэн системийн тайлан мэдээнээс өгөгдлөө цуглуулан MS-Office Excel програм дээр судалгаа шинжилгээ хийж үр дүн гаргадаг. Ингэж үр дүн гаргах нь их өгөгдөлд дүн шинжилгээ хийхэд ихээхэн цаг хугацаа, хүн хүч шаарддаг байна.

HeavenPro даатгалын үндсэн систем дээр өгөгдлийн тандалтын алгоритм болон арга техникүүдийг хөгжүүлэн нэвтрүүлж өгснөөр даатгалын байгууллага нь өөрийн ашиглаж буй үндсэн системдээ өгөгдлийн тандалтын аргуудыг ашиглан их өгөгдөлд дүн шинжилгээ хийх боломжийг нээн өгч байна.

Уг ажлын үр дүнд их өгөгдөлд дүн шинжилгээ хийх хүн хүч, цаг хугацааны хувьд хэмнэлт хийх, тухайн мэргэжилтнээс үл хамаарах шийдвэр гаргах, системийн өгөгдөлд системээс шууд дүн шинжилгээ хийж бодит үр дүнг гарган даатгалын байгууллагын цаашдын үйл ажиллагаанд ашиглах боломж бүрдэж байна.

Бидний судалгааны ажлаар сонгон авсан ХХХ даатгалын байгууллагын Жолоочийн хариуцлагын албан журмын даатгалын нөхөн төлбөрийн өгөгдөлд өгөгдлийн тандалт хийснээр дараах үр дүнг гаргасан.

УЛААНБААТАР хотын БАЯНЗҮРХ дүүрэгт 19.00-24.00 цагийн хооронд TOYOTA PRUIS маркын В ангиллын авто тээврийн хэрэгслийг 25-35 насны эрэгтэй жолооч нар замын хөдөлгөөнд оролцох үедээ

ЗАМЫН ХӨДӨЛГӨӨНД ОРОЛЦОХ ҮЕИЙН ЭРСДЭЛ-ийг их гаргадаг.

#### СУДАЛГААНЫ ИРЭЭДҮЙД ХИЙХ АЖИЛ

Судалгааны дараа дараагийн шатанд HeavenPro даатгалын үндсэн системд өгөгдлийн тандалтын бусад арга техникүүдийг судлан нэвтрүүлэх ба даатгалын байгууллагын харилцагчийг ангилж бизнесийн байгууллагын ойрын болон хэтийн төлөвт ашиглагдах шинэ мэдлэг, мэдээлэл гарган авахад чиглэгдэнэ.

#### АШИГЛАСАН МАТЕРИАЛ

- [1] Lucas Lau and Arun Tripathi, Mine Your Business-A Novel Application of Association Rules for Insurance Claims Analytics, Casualty Actuarial E-Forum, Winter 2011.
- [2] Lijia Guo, Applying Data Mining Techniques in Property/Casualty Insurance, Forums of the Casualty Actuarial Society, 2003
- [3] Амарбаясгалан Цацрал, Оюун-Эрдэнэ Намсрай, Их хэмжээний өгөгдөл дээр дүн шинжилгээ хийх боломжтой бизнесийн мэдээллийн системийг хэрэгжүүлэх аргачлал (Дэлгүүрийн бизнесийн мэдээллийн систем), Улаанбаатар, ММТ, 2015.
- [4] Шаравдорж Цэнд-Аюуш, “Өгөгдлийн уурхайг хөлбөмбөгийн тааварт ашиглах нь”, 2015, МУИС. ХШУИС
- [5] K. Umamaheswari, Dr. S. Janakiraman, Role of Data mining in Insurance Industry. COMPUSOFT, An international journal of advanced computer technology, 3 (6), June-2014
- [6] H. Lookman Sithic, T. Balasubramanian, Survey of Insurance Fraud Detection Using Data Mining Techniques, International Journal of Innovative Technology and Exploring Engineering (IJITEE), ISSN: 2278-3075, Volume-2, Issue-3, February 2013
- [7] Fawzi Elias Bekri, Dr. A. Govardhan, Association of Data Mining and Healthcare Domain: Issues and Current State of the Art, Global Journal of Computer Science and Technology, Volume 11 Issue 21 Version 1.0 December 2011, ISSN: 0975-4350
- [8] Abdhesh Gupta, Anwiti Jain, Life Insurance Recommender System Based on Association Rule Mining and Dual Clustering Method for Solving Cold-Start Problem, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 10, October 2013 ISSN: 2277 128X
- [9] A. B. Devale, Dr. R. V. Kulkarni, Applications of data mining techniques in life insurance, International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.2, No.4, July 2012