Оюутны сэтгэл ханамжийн үнэлгээний кластер шинжилгээ

Ц.Мөнгөнтуул[1], Л.Пүрэвдолгор[1], Ц.Жавзмаа[1], Т.Ахыт[1], П.Жаргалбат[1], Ч.Баасандарж[1], Ч.Галбадрах[1], Л.Ажнай[1]

[1]АШУҮИС, Био-Анагаахын сургууль, Физик, Мэдээлэл зүйн тэнхим
Цахим шуудан: purevdolgor@mnums.edu.mn, Утас: 95091107

Түлхүүр үг:

Асуумж судалгаа
Өгөгдлийн тандалт
k-means арга

Товч утга:

Үндэслэл: Их, дээд сургуулийн үйлчилгээний чанарын шалгуур үзүүлэлтийн нэг бол оюутны сэтгэл ханамж юм. **Зорилго:** Анагаахын Суурь боловсролын I дамжааны оюутнуудаас Анагаахын физикийн хичээлд хэр сэтгэл ханамжтай байгааг кластерын аргаар судлахаар зорилоо. **Арга, аргачлал:** Энэхүү зорилгын хүрээнд АШУҮИС-ийн Анагаахын суурь боловсролын I дамжааны анагаахын физик хичээлийг судалсан нийт 324 оюутнаас асуумж судалгаа авлаа. Сэтгэл ханамжийн түвшинг "*Маш Сайн*", "*Сайн*", "*Дунд* ", *Муу*", "*Маш Муу*" гэсэн таван түвшингээр тодорхойлов. Оюутнуудаас хичээлийн үнэлгээний 8, багшийн үнэлгээний 19, оюутны үнэлгээний 12 нийт 39 асуумжаар судалгаа авлаа. Судалгаанаас 39 талбартай, 324 бичлэгтэй мэдээллийн санг үүсгэв. Энэхүү мэдээллийн санд өгөгдөл тандалтын k-means аргаар MATLAB, R програмд кластер шинжилгээ хийв. **Үр дүн:** Өгөгдлийн санг H шалгуур (hopkins)-ыг бодоход 0.88, 0.87, 0.97 гэсэн утгууд гарлаа. Энэ нь кластер шинжилгээний аргуудыг ашиглах боломжтой болохыг харуулж байна. Хичээлийн үнэлгээний дэд санг гурван кластерт хуваалаа. Үүнд: I кластерт "*Сайн*" үнэлгээтэй 150 обьект 46.2%, II кластерт "*Дунд*" үнэлгээтэй 119 обьект 36.7%, III кластерт "*Муу*" үнэлгээтэй 54 обьект 16.6% байна. Багшийн үнэлгээний дэд санг гурван кластерт хуваахад II кластерт "*Сайн*" үнэлгээтэй 179 обьект 55.2%, III кластерт "*Дунд* " үнэлгээтэй 108 обьект 33.3% нь, I кластерт "*Муу*" үнэлгээтэй 36 обьект 11.1% байна. Оюутаны үнэлгээний дэд санг хоёр кластерт хуваахад II кластерт "*маш Сайн*" үнэлгээтэй 215 обьект 66.3%, I кластерт "*Муу*" үнэлгээтэй 108 обьект 33.3% байна. Энэхүү үүссэн кластеруудаа дүрслэлийн коэффициентоор (silhouette coefficient) үнэлэхэд нь хичээлийн үнэлгээний кластер 0.32, багшийн үнэлгээний кластер 0.31, оюутаны үнэлгээний 0.30 гарсан нь статистик ач холбогдолтой болохыг харуулж байна. **Дүгнэлт:** Эцэст нь дүгнэхэд кластер шинжилгээгээр Анагаахын физикийн хичээлийн загварт *"Сайн"* үнэлгээ 46.2%, *"Дунд"* үнэлгээ 36.7%, *"Муу"* үнэлгээ 16.6%, багшийн үнэлгээтэй загварт *"Сайн"* үнэлгээ 55.2%, *"Дунд"* үнэлгээ 33.3%, *"Муу"* үнэлгээ 11.1%, оюутны үнэлгээний загварт *"Сайн"* үнэлгээ 66.3%, *"Муу"* үнэлгээ 33.3% байна.

Үндэслэл: Боловсрол нь нийгмийн хөгжилд түлхэц үзүүлдэг гол хүч бөгөөд энэ салбарын хөгжилд багагүй хөрөнгө зарцуулсаар байгаа баримт байна. Зарцуулж буй хөрөнгийн үр ашгийг нэмэгдүүлэхэд чанарын асуудал тавигдаж, боловсролыг хэрхэн чанартай, илүү үр ашигтай болгох, түүний судалгааг хөгжүүлэх талаар боловсролын бүх түвшинд чанарын үнэлгээ хийсээр байна. Иймээс оюутнууд сургуулиас явуулж буй сургалт, багшийн харьцаа, хандлага ур чадварт хэр сэтгэл ханамжтай байгааг судлахаар зорилоо.

Шинжлэх ухааны олон салбарт цугларсан их хэмжээний мэдээллийг боловсруулах, шинжлэх, зүй тогтлыг илрүүлэн загварчлах, дүгнэх, урьдчилан таамаглах шинэ арга технологийг хөгжүүлэх зайлшгүй шаардлагатай. Өгөгдлийн тандалт нь их хэмжээний өгөгдөл дотроос далд зүй тогтол, утгын уялдаа холбоог

нээн илрүүлэх процесс бөгөөд өгөгдлийн сангийн систем, математик, статистик болон загварчилах аргуудаар мэдлэг олж авах аргачлал юм. Энгийнээр өгөгдлийн тандалт нь өгөгдлийн сангаас хамгийн хэрэгцээтэй мэдлэг нээх арга юм. Энэ судалгааны чиглэлийг "Data mining" буюу "Өгөгдлийн тандалт" юм [1-3] гэж судлаачид тодорхойлсон байна.

Өгөгдлийн тандалтад цугларсан их хэмжээний өгөгдлүүдийн хоорондын ерөнхий шинжийг тодорхойлох, өгөгдлүүдийн холбоо, хамааралд дүн шинжилгээ хийх, өгөгдлүүдийг нийтлэг шинжүүдээр нь бүлэглэх, тодорхой үзүүлэлтүүдээр урьдчилан таамаглах, хэт ялгаатай өгөгдлүүдийг ялгах гэх мэт тооцоолол хийхэд ангилалт (Classification) болон бүлэглэлтийн (Clustering) аргуудыг ашигладаг.

Зорилго: АШУҮИС-ийн Анагаахын суурь боловсролын I дамжааны анагаахын физик хичээлийг судалсан оюутнуудаас сэтгэл ханамжийн судалгаа авч, кластер шинжилгээ хийх.

Арга, аргачлал: Энэхүү зорилгын хүрээнд АШУҮИС-ийн Анагаахын суурь боловсролын I дамжааны анагаахын физик хичээлийг судалсан нийт 324 оюутнаас асуумж судалгаа авлаа. Сэтгэл ханамжийн түвшинг "*Маш Сайн*", "*Сайн*", "*Дунд*", *Муу*", "*Маш Муу*" гэсэн таван түвшингээр тодорхойлов. Оюутнуудаас хичээлийн үнэлгээний 8, багшийн үнэлгээний 19, оюутны үнэлгээний 12 нийт 39 асуумжаар судалгаа авлаа. Судалгаанаас 39 талбартай, 324 бичлэгтэй мэдээллийн санг үүсгэв. Энэхүү мэдээллийн санд өгөгдөл тандалтын k-means аргаар MATLAB, R програмд кластер шинжилгээ хийв.

Кластер шинжилгээ: Бүлэг дэх объектууд өөр хоорондоо адил, бусад бүлгийн объектоос өөр байхаар багцын объектуудыг хувааж, байрлуулахыг кластер шинжилгээ (cluster analysis) гэнэ. Энэхүү k - means аргад n объектыг агуулсан D багцтай гэж үзнэ. n объектыг $C_1 \dots C_k$ гэх мэт k кластерт $C_i \subset D$ ба $C_i \cap C_j = 0$ нөхцөлтөйгөөр хувааж байрлуулна. Хувааж байрлуулахдаа кластер бүрийн хувьд тухайн объектоос объект хүртэлх зайг квадрат алдааны нийлбэрээр тооцно[4]. Дүрслэлийн коэффициентоор (silhouette coefficient) үүссэн кластерыг үнэлнэ[3]. Дүрслэлийн коэффициентийг доорхи томьёогоор тооцов.

$$s(i) = \frac{(b(i) - a(i))}{max\{a(i), b(i)\}}$$

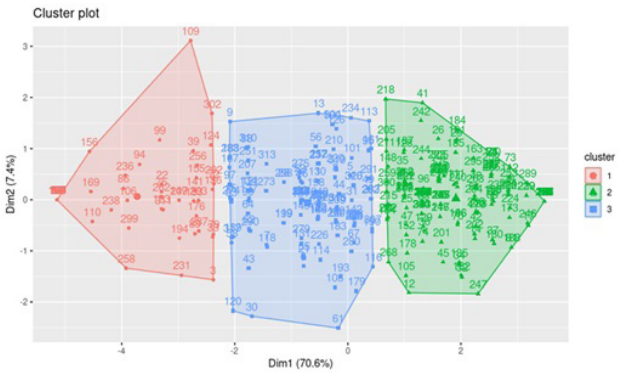
Дээрх тэгшитгэлд *a(i)* - *i* объектыг агуулж байгаа кластерын объект хоорондын дундаж зай (кластер доторх зай), *b(i)* - *i* объектыг агуулж кластерын объект ба өөр кластерийн объект хоорондын хамгийн бага зай болно[4,6-8]. Дүрслэлийн коэффициент нь - 1 < *s(i)* < 1 хооронд байна. Дүрслэлийн коэффициентын утгаас хамааран хүчтэй, зохистой, сул бүтэц байгаа эсэхийг ирэрхийлэх бөгөөд 0.26 -аас дээш утга авсан тохиолд олдсон бүтэц тохиромжтой гэж үздэг[5-6].

Үр дүн: Анагаахын суурь боловсролын I дамжааны анагаахын физик хичээлийг судалсан 324 оюутнаас 39 асуумжаар судалгаа авлаа. Энэхүү асуумж судалгаанаас 39 талбартай, 324 бичлэгтэй мэдээллийн санг үүсгэв. Үүссэн мэдээллийн сандаа MATLAB болон R програмд -means аргаар кластер шинжилгээ хийв[2,3]. Сэтгэл ханамжийн түвшинг сэтгэл ханамж "*Маш Сайн*", "*Сайн*", "*Дунд*", "*Муу*", "*Маш Муу*" гэсэн таван түвшингээр тодорхойлсон болно.

Оюутнуудаас гурван бүлгээр асуумжаар мэдээллийн сан бүрдүүлэн, MATLAB, R программын *k*-means аргуудаар бүлгүүдэд ижил төстэй шинжээр хувааж, *k*-means кластерын аргаар загварууд үүсгэлээ. Энэхүү загваруудын хуваагдсан кластерын тоо, кластер бүрийн дүрслэлийн коэффициент, объектуудын тоо, сэтгэл ханамжийн түвшинг *k*-means аргаар тооцож, загварын үнэлгээг Хүснэгт 1-д тус үзүүлэв. Үүссэн кластеруудын дүрслэлийн коэффицент *s(i)* нь 0.19-0.4 хооронд байгаа нь тохиромжтой ***Дундаж*** бүтэц олдсоныг илтгэж байна[4, 12].

Хичээлийг үнэлэх загвар: Хичээлийн үнэлэх санг *k*-means аргаар гурван кластерт хуваасныг дүрслэв. Үүнд сэтгэл ханамжийн түвшин *"Сайн"* 150 объекттэй II кластер, сэтгэл ханамжийн түвшин *"Дунд"* 119 объекттэй III кластер, сэтгэл ханамжийн түвшин *"Муу"* 54 объекттэй I кластерыг илэрхийлнэ (Зураг 1).

Багшийн оролцоог үнэлэх загвар: Хичээлд багшийн оролцоо, харьцаа хандлыгыг үнэлэх мэдээллийн санг гурван кластерт хуваагдсан загварыг Зураг 2-д харуулав.



Зураг 1. Хичээлийг үнэлэх загвар

Хүснэгт 1. Загварын үнэлгээ

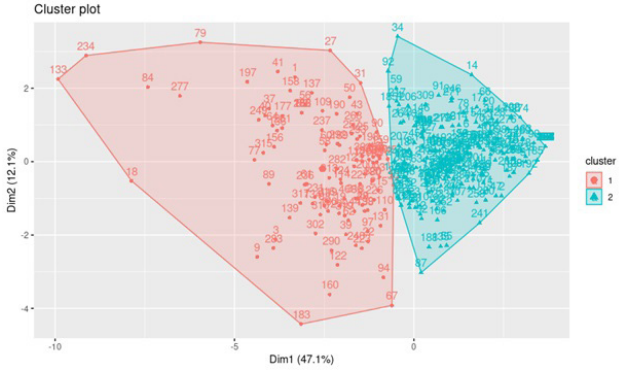
Хичээлийг үнэлэх загвар					Багшийн оролцоог үнэлэх загвар					Оюутны оролцоог үнэлэх загвар				
Кластерын дугаар	Объектын тоо	Объектын эзлэх хувь	коэффициент	Сэтгэл ханамжийн түвшин	**Кластерын дугаар**	Объектын тоо	Объектын эзлэх хувь	Дүрсийн коэффициент	Сэтгэл ханамжийн түвшин	**Кластерын дугаар**	Объектын тоо	Объектын эзлэх хувь	коэффициент	Сэтгэл ханамжийн түвшин
1	54	16.6	0.34	Муу	**1**	36	11.1	0.40	Муу	**1**	108	33.3	0.23	Муу
2	150	46.2	0.40	Сайн	**2**	179	55.2	0.40	Сайн	**2**	215	66.3	0.38	Сайн
3	119	36.7	0.24	Дунд	**3**	108	33.3	0.19	Дунд					

Үүнд сэтгэл ханамжийн түвшин *"Сайн"* 179 объекттэй II кластер, сэтгэл ханамжийн түвшин *"Дунд"* 108 объекттэй III кластер, сэтгэл ханамжийн түвшин *"Муу"* 36 объекттэй I кластерыг илэрхийлнэ (Зураг 2).



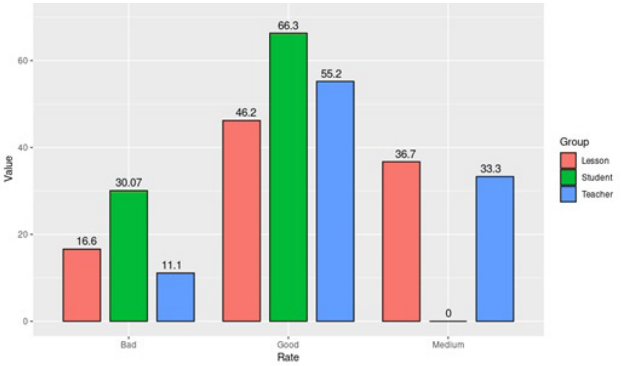
Зураг 2. Багшийн оролцоог үнэлэх загвар

Оюутныг үнэлэх загвар: Хоёр кластерт хуваагдсан оюутныг үнэлэх загварыг Зураг 3-т харуулав. Үүнд сэтгэл ханамжийн түвшин "Сайн" 215 объекттэй II кластер, сэтгэл ханамжийн түвшин "Муу" 108 объекттэй I кластерыг илэрхийлнэ (Зураг 3).



Зураг 3. Оюутныг үнэлэх загвар

Хичээл, багшийн оролцоо, оюутныг үнэлэх загваруудын үр дүнг нэгтгэв (Зураг 4).



Зураг 4. Сэтгэл ханамжийн үнэлгээний загваруудын дүн

Анагаахын физик хичээлийн сэтгэл ханамжийн түвшин 5% "Муу", 61% "Сайн", 33% "Дунд", багшийн оролцоо 3% "Муу", 69% "Сайн", 27% "Дунд", оюутнуудын оролцоо, идэвх 39% "Сайн", 29% "Дунд", 31% Маш сайн гэж үнэлсэн байна.

Хэлцэмж: 2017 онд хийсэн судалгаагаар шийдвэрийн модны загварт 32% нь сэтгэл ханамж "Маш сайн" болон "Сайн", 52% сэтгэл ханамж "Дунд", 16% сэтгэл ханамж "Муу" гэж үнэлэгдсэн байна[13]. Энэхүү судалгаатай харьцуулахад "Сайн", "Дунд", "Муу" үнэлгээний эзлэх хувь ойролцоо байна. Харин 2021 онд оюутнуудын цахим хичээлийн орчин, технологийн хэрэгцээ, оюутны эзэмшсэн мэдлэгийг кластер аргаар судлахад цахим сургалтын орчны загварт "Сайн" үгэлгээ 54.7% дунд үнэлгээ 45.3%, цахим даалгаварын гүйцэтгэл загварт "Сайн" үгэлгээ 61.3%, "Маш сайн" үнэлгээ 24.6%, "Дунд" үнэлгээ 14.1%, цахим сургалтаар эзэмшсэн мэдлэгийн үнэлгээ загварт "Сайн" үнэлгээ 45.9%, "Дунд" үнэлгээ 35.2%, "Маш сайн" үнэлгээ 18.9%, цахим сургалтын үр дүн загварт 81.7%, "Дунд" үнэлгээ 18.3% гарсан байна[14] . Энэ судалгаатай харьцуулахад "Сайн", "Дунд", "Муу" үнэлгээний эзлэх хувь бага зэрэг буурсан байна.

Дүгнэлт:

1. Анагаахын суурь боловсролын I дамжааны 324 оюутнаас анагаахын физикийн хичээлээр хичээлийг үнэлэх, багшийн оролцоог үнэлэх,

оюутны үнэлэх зэрэг бүлгээр асуумжаар өгөгдлийн сан үүсгэн, k-means аргаар ижил төстэй шинжээр бүлгүүдэд хуваааж, загварууд үүсгэлээ.
2. Хичээлийг үнэлэх загварын дүрслэлийн коэффициент 0.316, багшийн оролцоог үнэлэх загварын дүрслэлийн коэффициент 0.313 байгаа нь ач холбогдолтой бүтэц олдсоныг илтгэж байна.
3. Хичээлийг үнэлэх асуумжаар физикийн хичээлд хичээлийг үнэлэх загварт 46.2% *"Сайн"*, багшийн үнэлгээний загварт 55.2% *"Сайн"*, оюутны идэвх, оролцооны загварт 66.3 *"Сайн"* гарсан нь энэхүү сургалт амжилттай явагдсаныг илтгэж байна.

Ном зүй:

1. *Д. Сарангэрэл, "Өгөгдлийн тандалт болон үүлэн тооцооллын зарим арга зүйг эрүүл мэндийн салбарт нэвтрүүлэх судалгаа," Техникийн ш/у-ны д-рын (Ph.D) зэрэг горилсон диссертац, ШУТИС, УБ, 2013.*
2. *D. J. Hand, "Principles of data mining," Drug safety, vol. 30, no. 7, pp. 621-622, 2007.*
3. *I. H Witten, "Data mining," 2016.*
4. *J. Han, J. Pei, and M. Kamber, Data mining: concepts and techniques. Elsevier, 2011.*
5. *P.-N. Tan, M. Steinbach, and V. Kumar, Introduction to Data Mining. 2006.04.25.*
6. *J. A. Hartigan and M. A. Wong, "Algorithm AS 136: A k-means clustering algorithm," Journal of the royal statistical society. series c (applied statistics), vol. 28, no. 1, pp. 100-108, 1979.*
7. *A. Likas, N. Vlassis, and J. J. Verbeek, "The global k-means clustering algorithm," Pattern recognition, vol. 36, no. 2, pp. 451-461, 2003.*
8. *T. d. science, "K-means Clustering: Algorithm, Applications, Evaluation Methods, and Drawbacks."*
9. *K. Sasirekha and P. Baby, "Agglomerative hierarchical clustering algorithm-a," International Journal of Scientific and Research Publications, vol. 83, p. 83, 2013.*
10. *F. Murtagh and P. Contreras, "Algorithms for hierarchical clustering: an overview," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 2, no. 1, pp. 86-97, 2012.*
11. *L. Kaufman and P. J. Rousseeuw, Finding groups in data: an introduction to cluster analysis. John Wiley & Sons, 2009.*
12. *RDocumentation. Visual Enhancement Of Clustering Analysis. Available: https://www.rdocumentation.org/packages/factoextra/versions/1.0.5/topics/eclust*

Cluster analysis of student satisfaction

Munguntuu Tsl[1], Purevdolgor L[1], Javzmaa Ts[1], Akhyt T[1], Jargalbat P[1], Baasandarj Ch[1], Galbadrach Ch[1], Ajnai L[1]

[1]Department of Physics and Informatics, School of Biomedicine, MNUMS
E-mail: purevdolgor@mnums.edu.mn, Tel: +976-95091107

Background: One of the indicators of the quality of university services is student satisfaction.
Aim: We aimed to study the level of satisfaction of students in the first year of premedical courses in the course of Medical Physics using the cluster method.
Materials and Methods: In the framework of this goal, a questionnaire was collected from a total of 324 students who studied the medical physics course of the 1st course of the premedical course of MNUMS. When determining the level of satisfaction, the answers were obtained on five levels of satisfaction: "Excellent", "Good", "Medium", "Bad" and "Very bad". A total of 39 questionnaires were collected from students, 8 for course evaluation, 19 for teacher evaluation, and 12 for student evaluation. From the research, a database with 39 fields and 324 records was created.
Results: In this database, cluster analysis was performed in MATLAB and R programs using the k-means method of data mining. Calculated the Hopkins statistic in the created database, the values are 0.88, 0.87, and 0.97. This shows that cluster analysis methods can be used. The course evaluation sub-fund is divided into three clusters. Among them: cluster I has 150 objects with a "Good" rating of 46.2%, cluster II has 119 objects with a "medium" rating of 36.7%, and Cluster III has 54 objects with a "Good" rating of 16.6%.
The teacher evaluation sub-base into three clusters, there are 179 objects with a "good" rating of 55.2% in cluster II, 108 objects with an "Average" rating of 33.3% in cluster III, and 36 objects with an "excellent" rating in cluster I of 11.1%. The sub-base of student evaluations into two clusters, cluster II has 215 objects with an "excellent" rating of 66.3%, and cluster I have 108 objects with an "excellent" rating of 33.3%. Evaluating the resulting clusters with the Silhouette coefficient, 0.32 for the course evaluation cluster, 0.31 for the teacher evaluation cluster, and 0.30 for student evaluation show statistical significance.
Conclusion: Finally, to conclude, cluster analysis in the model of the medical physics lesson "Good"-46.2%, "Middle"-36.7%, "bad"-16.6%; 55.2% -"Good", 33.3% - "Middle", 11.1%-"Bad" in the teacher evaluation model; 66.3%-"Good" and 33.3% of "bad" in the student evaluation model.

Keywords: Questionnaire, Data mining, k-means method