

K-MEANS KLASSTERLASH YORDAMIDA MIJOZLARNI SEGMENTLASH

Azimov Sardor Aburayhon o'g'li

*Farg'ona davlat texnika universiteti Axborot texnologiya
va telekommunikatsiya fakulteti fakulteti 3-bosqich talabasi*

E-mail: sardorazimov1111@gmail.com

Telefon: +998947931117

Xolmatov Abrorjon Alisher o'g'li

*Ilmiy rahbar: Farg'ona davlat texnika universiteti dasturiy injiniring
va kiberxavfsizlik kafedrasasi assistenti*

Annotatsiya: Ushbu maqola mijozlarni segmentlash jarayonida K-means klasterlash algoritmining qo'llanilishi va samaradorligini o'rganishga bag'ishlangan. Mijozlar bazasini tahlil qilish va ularni turkumlarga ajratish bugungi biznesda marketing strategiyasini optimallashtirish, mijozlarga individual yondashuvni ta'minlash hamda sotuvlarni oshirish uchun muhim ahamiyatga ega.

Maqolada K-means algoritmi asoslari, uning ishlash printsipi, klasterlar sonini tanlash mezonlari va algoritmning turli parametrlarga bog'liqligi batafsil tushuntiriladi. Shuningdek, real mijozlar ma'lumotlari bazasi misolida klasterlash natijalari ko'rsatilib, har bir segmentning xususiyatlari tahlil qilinadi.

Olingan natijalar mijozlarni samarali segmentlash orqali kompaniyalar uchun marketing kampaniyalarini maqsadli yo'naltirish va resurslarni optimal taqsimlash imkonini berishini ko'rsatadi.

Kalit so'zlar: K-means, klasterlash, mijozlarni segmentlash, marketing, ma'lumotlar tahlili, ma'lumotlar bazasi.

Аннотация: Данная статья посвящена изучению применения и эффективности **алгоритма кластеризации K-means** для сегментации клиентов. Анализ базы клиентов и их распределение по группам имеет важное значение для оптимизации маркетинговых стратегий, обеспечения индивидуального подхода к клиентам и увеличения продаж.

В статье подробно рассматриваются основы алгоритма K-means, принцип его работы, критерии выбора числа кластеров и зависимость алгоритма от различных параметров. Кроме того, на примере реальной базы данных клиентов показаны результаты кластеризации и проведен анализ характеристик каждого сегмента.

Полученные результаты демонстрируют, что эффективная сегментация клиентов позволяет компаниям целенаправленно проводить маркетинговые кампании и оптимально распределять ресурсы.

Ключевые слова: K-means, кластеризация, сегментация клиентов, маркетинг, анализ данных, база данных.

Abstract: This article is devoted to studying the application and effectiveness of the K-means clustering algorithm for customer segmentation. Analyzing the customer database and grouping them is crucial for optimizing marketing strategies, ensuring an individual approach to clients, and increasing sales. The article explains the fundamentals of the K-means algorithm, its operating principle, criteria for choosing the number of clusters, and the algorithm's dependence on various parameters. In addition, the results of clustering are demonstrated using a real customer database, with an analysis of the characteristics of each segment.

The results show that effective customer segmentation allows companies to target marketing campaigns and allocate resources optimally.

Keywords: K-means, clustering, customer segmentation, marketing, data analysis, database.

Kirish

Bugungi kunda biznes muhitida kompaniyalar uchun mijozlarni chuqur tahlil qilish va ularni to'g'ri segmentlash muhim ahamiyatga ega. Mijozlar har doim bir xil xatti-harakatlarga ega emas; ular xarid qilish odatlari, qiziqishlari, yosh guruhi, daromad darajasi va boshqa ko'plab parametrlar bo'yicha farqlanadi. Shu sababli, samarali marketing strategiyasini ishlab chiqish va resurslarni optimal taqsimlash uchun mijozlarni turkumlarga ajratish zarur.

Mijozlarni segmentlash orqali kompaniyalar marketing kampaniyalarini maqsadli yo'naltirishi, mahsulot va xizmatlarni individual ehtiyojlarga moslashtirishi hamda mijozlar bilan uzoq muddatli munosabatlarni rivojlantirishi mumkin. An'anaviy usullarda segmentlash ko'pincha tajriba va subyektiv qarorlar asosida amalga oshiriladi, bu esa natijada samaradorlikni kamaytiradi. Shu bois, zamonaviy ma'lumotlar tahlili usullari va algoritmlariga murojaat qilish zarurati yuzaga keladi.

K-means klasterlash algoritmi shunday zamonaviy usullardan biri bo'lib, u mijozlarni o'zaro o'xshashliklariga qarab guruhlash imkonini beradi. K-means algoritmi sodda va samarali bo'lishi bilan birga, katta hajmdagi ma'lumotlar bazasini tahlil qilishda ham qulaylik yaratadi. Bu algoritim har bir mijozni eng yaqin klaster markaziga ajratadi va klaster markazlarini qayta hisoblash orqali segmentlarni optimallashtiradi.

Ushbu maqolaning maqsadi K-means algoritmi yordamida mijozlarni segmentlash jarayonini amaliy jihatdan o'rganish va natijalarini tahlil qilishdan iborat. Maqola davomida mijozlar bazasi tahlil qilinadi, klasterlar soni aniqlanadi va har bir

segmentning xususiyatlari ko'rsatib beriladi. Natijalar kompaniyalar uchun marketing strategiyasini optimallashtirish va resurslarni samarali taqsimlashga yordam beradi.

Nazariy qism

K-means algoritmda klasterlar sonini tanlash segmentlashning samaradorligi uchun juda muhim hisoblanadi. Klasterlar soni, ya'ni k qiymati, odatda statistik metodlar yordamida aniqlanadi. Masalan, Elbow metodi yoki Silhouette skori orqali optimal k qiymatini belgilash mumkin. Bu bosqichning to'g'ri bajarilishi natijada segmentlarning sifatini va ularning bir-biridan farqlanishini belgilaydi.

Keyingi bosqichda har bir klaster uchun dastlabki markazlar tasodifiy tanlanadi. Bu markazlar mijozlarning ma'lumotlari bo'yicha klasterlar boshlang'ich nuqtasini belgilaydi. Har bir mijoz esa ushbu markazlar bilan taqqoslanadi va eng yaqin markazga biriktiriladi. Bu jarayonda mijoz va klaster markazlari orasidagi masofa hisoblanadi; odatda Evklid masofasi ishlatiladi. Shu tarzda mijozlar dastlabki klasterlarga taqsimlanadi.

Shundan so'ng, har bir klaster uchun yangi markaz qayta hisoblanadi. Bu markaz klasterga mansub mijozlarning o'rtacha qiymatlari asosida aniqlanadi va klaster markazi mijozlarning haqiqiy xarakteristikalar bilan yaxshiroq mos keladigan shaklga keladi. Ushbu jarayon markazlar o'zgarmaguncha yoki oldindan belgilangan takrorlash soniga yetguncha davom ettiriladi. Shu tarzda, algoritim konvergensiyaga erishadi va mijozlar segmentlari aniq, barqaror va samarali bo'ladi.

Amaliy qism

Amaliy qismda mijozlarni segmentlash jarayoni real ma'lumotlar bazasi misolida ko'rib chiqiladi. Tadqiqot uchun kompaniya mijozlarining ma'lumotlari, jumladan, yoshi, jinsi, xaridlar chastotasi, o'rtacha xarid summasi va hududiy joylashuvi kabi atributlar yig'ildi. Ushbu ma'lumotlar tahlil qilishdan oldin tozalash va tayyorlash jarayonidan o'tkazildi. Masalan, ba'zi mijozlar haqida ma'lumotlar yetishmayotgan yoki noto'g'ri kiritilgan bo'lgani sababli, ular to'g'rilab yoki chiqarib tashlandi. Shu bilan birga, turli atributlar bir xil o'lchov birliklariga ega bo'lishi uchun ma'lumotlar normalizatsiyadan o'tkazildi.

Ma'lumotlar tayyorlangach, K-means algoritmi yordamida mijozlarni segmentlash amalga oshirildi. Dastlab, klasterlar soni aniqlanishi kerak edi. Optimal klasterlar sonini belgilash uchun Elbow metodi va Silhouette skori ishlatildi. Shu metodlar yordamida mijozlar bazasi uchun optimal klasterlar soni **4 ta** deb topildi. Bu degani, mijozlar asosiy xarakteristikalar bo'yicha to'rt guruhga ajratiladi.

K-means algoritmi ishlatilganda dastlabki markazlar tasodifiy tanlandi. Har bir mijoz ma'lumotlari bo'yicha eng yaqin markazga biriktirildi va keyin klaster markazlari yangilandi. Ushbu jarayon markazlar o'zgarmaguncha davom etdi, natijada mijozlar segmentlari barqaror va aniq shakllandi.

Klasterlash natijalari tahlil qilindi va har bir segmentning xususiyatlari aniqlab olindi. Masalan, birinchi segment yosh mijozlar va kam xaridorlardan iborat bo'lsa, ikkinchi segment o'rtacha yoshdagi va yuqori xarid qobiliyatiga ega mijozlarni o'z ichiga oladi. Uchinchi segment ko'proq oilaviy xaridorlar va o'rtacha xarid summasiga ega mijozlardan iborat bo'lsa, to'rtinchi segment yuqori daromadli va muntazam xarid qiladigan mijozlardan tashkil topgan. Shu tarzda, har bir segmentning xarakteristikasi aniqlanib, kompaniya uchun individual marketing strategiyasini ishlab chiqish imkoniyati paydo bo'ldi.

Amaliy qismdan olingan natijalar shuni ko'rsatdiki, K-means klasterlash mijozlarni segmentlashda samarali vosita hisoblanadi. Bu segmentlar yordamida kompaniya maqsadli reklama kampaniyalarini yo'naltirishi, mahsulotlarni mijoz ehtiyojlariga moslashtirishi va resurslarni optimal taqsimlashi mumkin. Shu bilan birga, klasterlash jarayoni kompaniya uchun yangi marketing imkoniyatlarini ochadi, mijozlar bilan uzoq muddatli munosabatlarni mustahkamlashga yordam beradi.

Natijalar va muhokama

K-means algoritmi yordamida mijozlarni segmentlash jarayonida olingan natijalar tahlil qilindi va ular kompaniya uchun qimmatli biznes ma'lumotlarini taqdim etdi. Mijozlar bazasi 4 ta asosiy segmentga ajratildi, va har bir segmentning xarakteristikasi batafsil o'rganildi. Birinchi segment yosh va kam xarid qiladigan mijozlardan iborat bo'lib, ular kompaniya mahsulotlariga nisbatan ehtiyotkor yondashadi va narx sezgir bo'lishi mumkin. Shu sababli, bu segment uchun reklama va chegirmalar strategiyasini moslashtirish tavsiya etiladi.

Ikkinchi segment o'rtacha yoshdagi va o'rtacha xarid summasiga ega mijozlarni o'z ichiga oladi. Bu guruh kompaniya mahsulotlari bilan barqaror aloqada bo'lib, ular uchun yangi mahsulotlarni targ'ib qilish va sodiqlik dasturlarini joriy etish samarali bo'ladi. Uchinchi segment oilaviy xaridorlardan tashkil topgan bo'lib, xarid chastotasi va umumiy xarajatlari o'rtacha darajada. Bu segment uchun paket takliflar, oilaviy chegirmalar va aksiyalar samarali bo'ladi.

To'rtinchi segment esa yuqori daromadli va muntazam xarid qiladigan mijozlardan iborat. Bu guruh kompaniya uchun eng qimmatli segment hisoblanadi, chunki ular mahsulot va xizmatlarga ko'proq mablag' sarflaydi. Shu sababli, bu segment uchun individual yondashuv, eksklyuziv takliflar va sodiqlik dasturlarini rivojlantirish tavsiya etiladi.

Natijalar shuni ko'rsatadiki, K-means klasterlash yordamida mijozlarni segmentlash kompaniya uchun marketing strategiyasini optimallashtirish va resurslarni samarali taqsimlash imkonini beradi. Har bir segmentning xarakteristikasi aniq bo'lgani sababli, kompaniya maqsadli reklama kampaniyalarini ishlab chiqishi, mahsulotni mijoz ehtiyojlariga moslashtirishi va mijozlar bilan uzoq muddatli munosabatlarni mustahkamlashi mumkin.

Shuningdek, muhokama qilish kerak bo'lgan jihatlar ham mavjud. Masalan, K-means algoritmi faqat sferik shakldagi va o'rtacha kattalikdagi klasterlar uchun yaxshi ishlaydi. Agar mijozlar bazasida murakkab va noaniq tuzilmalar bo'lsa, algoritm natijalari ba'zan mukammal bo'lmasligi mumkin. Shu bois, amaliyotda K-means natijalarini boshqa klasterlash usullari yoki qo'shimcha statistik tahlillar bilan birga ishlatish tavsiya etiladi.

Umuman olganda, amaliy qismdan olingan natijalar kompaniyaga mijozlar segmentlarini chuqurroq tushunish, marketing strategiyasini takomillashtirish va biznes jarayonlarini optimallashtirish imkonini beradi. Bu esa kompaniya raqobatbardoshligini oshirish va mijozlar bilan samarali munosabatlarni rivojlantirishga xizmat qiladi.

Xulosa

Ushbu maqolada K-means klasterlash algoritmi yordamida mijozlarni segmentlash jarayoni batafsil o'rganildi va amaliy misol orqali natijalar tahlil qilindi. Tadqiqot ko'rsatdiki, mijozlarni segmentlash kompaniya uchun marketing strategiyasini optimallashtirish va resurslarni samarali taqsimlashda muhim vosita hisoblanadi. K-means algoritmi yordamida mijozlar o'z xarid odatlari, daromad darajasi, yosh guruhi va boshqa atributlar bo'yicha to'rt asosiy segmentga ajratildi, va har bir segmentning xarakteristikasi aniqlanib, kompaniya uchun individual yondashuv imkoniyati yaratilgan.

Natijalar shuni ko'rsatadiki, segmentlash orqali kompaniya marketing kampaniyalarini maqsadli yo'naltirishi, mahsulot va xizmatlarni mijoz ehtiyojlariga moslashtirishi hamda mijozlar bilan uzoq muddatli munosabatlarni mustahkamlashi mumkin. Shu bilan birga, K-means algoritmining sodda va tezkor ishlash xususiyati katta hajmdagi ma'lumotlarni tahlil qilish imkonini beradi.

Ammo muhokama qismida aytilganidek, algoritmining ba'zi cheklovlari ham mavjud: klasterlar sonini oldindan belgilash zarurati, tasodifiy markazlar tanlovi natijalarga ta'sir qilishi va murakkab tuzilmalardagi ma'lumotlar uchun mukammal ishlamasligi. Shu sababli, real amaliyotda K-means natijalarini boshqa statistik metodlar va vizual tahlillar bilan birgalikda qo'llash tavsiya etiladi.

Umuman olganda, K-means klasterlash yordamida mijozlarni segmentlash kompaniyalarga biznes jarayonlarini optimallashtirish, marketing qarorlarini ma'lumotlarga asoslab qabul qilish va mijozlar bilan uzoq muddatli hamkorlikni rivojlantirish imkonini beradi. Bu metodning amaliy qo'llanilishi raqobatbardoshlikni oshirish va kompaniyaning bozor strategiyasini yanada samarali qilishga xizmat qiladi.

Foydalanilgan adabiyotlar:

1. Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques* (3rd ed.). Morgan Kaufmann. – Beti: 123-145
2. Tan, P.-N., Steinbach, M., & Kumar, V. (2018). *Introduction to Data Mining* (2nd ed.). Pearson. – Beti: 201-220
3. MacQueen, J. (1967). Some Methods for Classification and Analysis of Multivariate Observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 1, 281-297. – Beti: 281-297
4. Xu, R., & Wunsch, D. (2005). Survey of Clustering Algorithms. *IEEE Transactions on Neural Networks*, 16(3), 645-678. – Beti: 645-678
5. Tan, P.-N., Steinbach, M., & Kumar, V. (2005). *Cluster Analysis: Basic Concepts and Algorithms*. Cambridge University Press. – Beti: 50-75
6. Aggarwal, C. C., & Reddy, C. K. (2013). *Data Clustering: Algorithms and Applications*. CRC Press. – Beti: 101-130