

ISODATA ALGORITMI VA UNING QO‘LLANILISHI

*Farg‘ona davlat universiteti amaliy matematika va
informatika kafedrasida dotsenti*

Onarqulov Maqsadjon Karimberdiyevich

maxmaqsad@gmail.com

*Farg‘ona davlat universiteti amaliy matematika va
informatika kafedrasida o‘qituvchisi*

Yusupov Mirsaid Abdulaziz o‘g‘li

mirsaidbeky@gmail.com

Farg‘ona davlat universiteti 3-kurs talabasi

Olimova Lobarxon Sherzod qizi

lobarxonkamolova0104@gmail.com

Annotatsiya: Ushbu maqolada ISODATA (Iterative Self-Organizing Data Analysis Technique) algoritmi va uning klasterlash jarayonidagi qo‘llanilishi tahlil qilinadi. ISODATA algoritmi K-o‘rtacha algoritmining rivojlangan shakli bo‘lib, ma‘lumotlarni avtomatik ravishda guruhlarga ajratish uchun ishlatiladi. Algoritmning ishlash mexanizmi, asosiy xususiyatlari, afzalliklari va kamchiliklari ko‘rib chiqiladi. Shuningdek, uning tasvirni qayta ishlash, geografik ma‘lumotlarni tahlil qilish, marketing, biologiya va boshqa sohalaridagi amaliy qo‘llanilish imkoniyatlari yoritib beriladi. Mazkur maqola ISODATA algoritmi haqida chuqurroq tushunchaga ega bo‘lishni istagan tadqiqotchilar va ma‘lumotlar tahlili bo‘yicha mutaxassislar uchun mo‘ljallangan.

Kalit so‘zlar: ISODATA, klasterlash, K-o‘rtacha, ma‘lumotlar tahlili, mashinaviy o‘rganish, iterativ algoritmlar, tasvirni qayta ishlash, geografik axborot tizimlari, marketing tahlili, biologik ma‘lumotlar.

Annotation: This article explores the ISODATA (Iterative Self-Organizing Data Analysis Technique) algorithm and its application in clustering processes.

ISODATA, an advanced version of the K-means algorithm, is used for automatically grouping data. The article examines the working mechanism of the algorithm, its key features, advantages, and limitations. Furthermore, its practical applications in image processing, geographic data analysis, marketing, biology, and other fields are highlighted. This article is intended for researchers and data analysis professionals seeking a deeper understanding of the ISODATA algorithm.

Keywords: *ISODATA, clustering, K-means, data analysis, machine learning, iterative algorithms, image processing, geographic information systems, marketing analysis, biological data.*

Аннотации: *В данной статье рассматривается алгоритм ISODATA (Iterative Self-Organizing Data Analysis Technique) и его применение в задачах кластеризации. ISODATA является усовершенствованной версией алгоритма K-средних, предназначенной для автоматического распределения данных на группы. В статье анализируется механизм работы алгоритма, его основные особенности, преимущества и ограничения. Также освещены практические приложения ISODATA в обработке изображений, анализе географических данных, маркетинге, биологии и других областях. Статья предназначена для исследователей и специалистов по анализу данных, желающих углубить свои знания об алгоритме ISODATA.*

Ключевые слова: *ISODATA, кластеризация, K-средние, анализ данных, машинное обучение, итеративные алгоритмы, обработка изображений, географические информационные системы, маркетинговый анализ, биологические данные.*

Kirish

Hozirgi davrda katta hajmdagi ma'lumotlarni samarali qayta ishlash va tahlil qilish dolzarb vazifalardan biriga aylangan. Bunda klasterlash algoritmlari muhim ahamiyat kasb etadi. Klasterlash — bu ma'lumotlarni o'zaro o'xshashlik darajasiga qarab guruhlariga ajratish jarayonidir. Mazkur usul tasniflashni talab qilmaydigan (nazoratsiz o'rganish) vazifalarni hal qilishda keng qo'llaniladi.

ISODATA (Iterative Self-Organizing Data Analysis Technique) algoritmi klasterlash algoritmlaridan biri bo‘lib, u K-o‘rtacha algoritmi asosida ishlab chiqilgan va yanada moslashuvchanligi bilan ajralib turadi. ISODATA algoritmi katta hajmdagi murakkab ma’lumotlarni qayta ishlashga mo‘ljallangan bo‘lib, u ma’lumotlar ichidagi yashirin tuzilmani aniqlashda samarali ishlaydi.

Ushbu maqolada ISODATA algoritmining ishlash jarayoni, uning afzallik va cheklovlari, shuningdek, turli sohalarda qo‘llanilish misollari muhokama qilinadi. Algoritmning moslashuvchanligi va samaradorligi uni katta hajmdagi ma’lumotlarni tahlil qilishda keng qo‘llaniladigan asosiy vositalardan biriga aylantiradi.

ISODATA (Iterative Self-Organizing Data Analysis Technique) algoritmi — bu klasterlash algoritmi bo‘lib, murakkab va katta hajmdagi ma’lumotlarni avtomatik ravishda guruhlariga ajratish uchun ishlatiladi. U K-o‘rtacha algoritmining yanada moslashuvchan shakli bo‘lib, ma’lumotlarni bir necha iteratsiya davomida optimallashtirishga asoslangan.

ISODATA Algoritmi Nima?

ISODATA — iterativ klasterlash algoritmi bo‘lib, quyidagi jarayonlar orqali ma’lumotlarni klasterlarga ajratadi:

Klasterlarni birlashtirish yoki bo‘lish.

Klaster markazlarini yangilash.

Ma’lumotlarning klasterlar o‘rtasida qayta taqsimlanishi.

Bu algoritm klasterlar sonini avtomatik ravishda aniqlashga harakat qiladi va bu jarayon davomida bir necha marta moslashuvchan tahlillarni amalga oshiradi.

ISODATA Algoritmi Ishlash Bosqichlari

1. Boshlang‘ich Klaster Markazlarini Tanlash:

Ma’lumotlar bo‘yicha dastlabki ta klaster markazlari tasodifiy tanlanadi.

2. Ma’lumotlarni Klasterlarga Ajratish:

Har bir nuqta o‘ziga yaqin klaster markaziga bog‘lanadi. Bu masofa odatda Evklid masofasi yordamida aniqlanadi.

3. Klaster Markazlarini Yangilash:

Har bir klaster uchun yangi markazlar hisoblanadi

4. Klasterlarni Birlashtirish yoki Bo‘lish:

Birlashtirish: Klasterlar orasidagi masofa oldindan belgilangan chegaradan kichik bo‘lsa, ular birlashtiriladi.

Bo‘lish: Klasterning dispersiyasi belgilangan chegaradan katta bo‘lsa, klaster ikkiga bo‘linadi.

5. Iteratsiyalarni takrorlash:

Yuqoridagi jarayon bir necha marta takrorlanadi, to klasterlar soni va joylashuvi barqarorlashguncha.

6. Natijani Baholash.

ISODATA Algoritmi Xususiyatlari

Klasterlar soni dastlabki holatda beriladi, lekin algoritm ularni birlashtirib yoki bo‘lib moslashadi.

Katta hajmdagi murakkab ma’lumotlarni samarali qayta ishlaydi.

Klasterlarning shakli va o‘lchamlari moslashuvchan.

ISODATA Algoritmining Qo‘llanilishi

1. Tasvirni Qayta Ishlash

ISODATA algoritmi tasvirlarni segmentatsiya qilishda keng qo‘llaniladi. Masalan, sun’iy yo‘ldosh rasmlarini tahlil qilishda ma’lumotlarni mintaqalarga ajratish.

2. Ma’lumotlarni Tahlil Qilish

Murakkab va katta hajmdagi ma’lumotlarni guruhlash orqali mantiqiy tuzilishni aniqlash. Bu ijtimoiy tadqiqotlar va biznes ma’lumotlarini tahlil qilishda foydali.

3. Geografik Tizimlar (GIS)

Geografik ma’lumotlarni, masalan, tuproq turlarini yoki iqlim zonalarini klasterlashda qo‘llaniladi.

4. Biologik Tahlillar

DNK yoki genetik ma'lumotlarni klasterlash orqali guruhlarga ajratish va yangi bog'liqliklarni aniqlash uchun ishlatiladi.

5. Marketing va Mijoz Tahlili

Mijozlarning xatti-harakatlarini guruhlash orqali mahsulotlarni maqsadli ravishda targ'ib qilish.

ISODATA Algoritmining Afzalliklari

1. Klasterlarning sonini oldindan aniqlashga ehtiyoj yo'q.
2. Klasterlarning shakli va o'lchamlari moslashuvchan bo'lib, algoritmi turli turdagi ma'lumotlarga moslashadi.
3. Birlashtirish va bo'lish jarayonlari murakkab klaster tuzilmalarini samarali qayta ishlaydi.

ISODATA Algoritmining Kamchiliklari

1. Parametrlarni (masalan, dispersiya chegarasi, masofa chegarasi) to'g'ri tanlash qiyin bo'lishi mumkin.
2. Hisoblash jihatidan qimmat (iteratsiyalar ko'p bo'lsa).
3. Ma'lumotlar kuchli shovqinli bo'lsa, natijalar noaniq bo'lishi mumkin.

Xulosa

ISODATA algoritmi o'zining moslashuvchanligi va samaradorligi bilan ko'plab tarmoqlarda muvaffaqiyatli qo'llanilmoqda. Uning iterativ ishlash mexanizmi va klasterlarni avtomatik ravishda optimallashtirish imkoniyati algoritmi katta hajmdagi murakkab ma'lumotlar bilan ishlashda muhim vosita sifatida namoyon qiladi. Parametrlarni to'g'ri tanlash orqali ISODATA algoritmi juda aniq va foydali natijalar bera oladi.

ISODATA algoritmi murakkab va katta hajmdagi ma'lumotlarni samarali klasterlash uchun moslashuvchan va qulay vositadir. Uning K-o'rtacha algoritmidan farqi — klasterlar sonini avtomatik ravishda o'zgartirish va ma'lumotlar ichidagi yashirin tuzilmani aniqlash qobiliyatidir. Tasvirni qayta ishlash, geografik axborot tizimlari, marketing, biologiya kabi turli sohalarda ISODATA algoritmi murakkab ma'lumotlar bilan ishlashda o'zining samaradorligini isbotlagan.

Shu bilan birga, algoritm parametrlarni to‘g‘ri tanlashni talab qiladi va hisoblash xarajatlari yuqoriligi sababli katta hajmdagi ma’lumotlar uchun optimallashtirishga ehtiyoj sezadi. ISODATA algoritmining imkoniyatlari va cheklovlarini to‘g‘ri tahlil qilish orqali uni ko‘plab ilmiy va amaliy vazifalarni yechishda samarali qo‘llash mumkin.

ISODATA algoritmi nafaqat ma’lumotlarni tahlil qilish uchun qulay vosita, balki zamonaviy ma’lumotlar ilmida muhim o‘ringa ega bo‘lgan nazoratsiz o‘rganish texnikasi hisoblanadi.

FOYDALANILGAN ADABIYOTLAR

Онаркулов, М. К. (2024). ИНТЕГРАЦИЯ НЕЙРОННЫХ СЕТЕЙ В РАСПОЗНАВАНИИ РЕЧИ И РАЗРАБОТКА НОВЫХ МЕТОДОВ. ОБРАЗОВАНИЕ НАУКА И ИННОВАЦИОННЫЕ ИДЕИ В МИРЕ, 47(7), 193-197.

Онаркулов, М. К., угли Юсупов, М. А., & угли Умиржонов, Л. А. (2023). ПРИМИНЕНИЕ НЕЙРОННЫХ СЕТЕЙ В РАСПОЗНАВАНИИ РЕЧИ. Educational Research in Universal Sciences, 2(3), 1206-1210.

Umarjon o‘g, Y. Y. L. (2024). PROBLEMS AND MODERN TRENDS IN COMPUTING ENGINEERING. Multidisciplinary and Multidimensional Journal, 3(6), 17-21.

Onarkulov, M., & Nabijonov, A. (2024). DB2: KATTA HAJMDAGI MA’LUMOTLARNI SAMARALI BOSHQARISH. Инновационные исследования в науке, 3(5), 99-104.

Onarkulov, M., & G‘oyibova, G. (2024). SQL SERVER ILOVASIDA UNIVERSAL MODELNI QO‘LLASH VA KONVERTATSIYA ETISHNING SAMARALI YO‘LLARI. Академические исследования в современной науке, 3(18), 147-152.

Karimberdiyevich, M. O. (2024). RBF TURLARINING O ‘QITISH ALGORITMI VA XOR MASALASI. worldly knowledge conferens, 8(1), 176-180.

Karimberdiyevich, M. O. (2024). GAUSS FUNKSIYASI. worldly knowledge conferens, 8(1), 239-244.

Karimberdiyevich, O. M. (2024). SQL TILIDA PROTSEDURA VA FUNKSIYALARDAN FOYDALANISHNING AHAMIYATI. worldly knowledge conferens, 8(1), 145-148.

Karimberdiyevich, O. M. (2024). BIR QATLAMLI NEYRON TO'RLARI VA ULARNI YARATISH USULLARI. IQRO INDEXING, 9(2), 104-108.

Onarkulov, M., & Omonaliyeva, E. (2024). QARORLAR DARAXTI VA UNI KIRITISH ALGORITIMI. Science and innovation in the education system, 3(6), 66-73.

Onarkulov, M., & Isaqova, S. (2024). NEYROCHIPLAR, MAXSUS MATRITSALI KUCHAYTIRGICHLAR VA NEYROEMULYATORLAR. Science and innovation in the education system, 3(6), 52-58.

Karimberdiyevich, M. O. (2024). EKSPERT TIZIMLARI YARATISH VA ULARNING MUAMMOLARI. ILM-FAN YANGILIKLARI KONFERENSIYASI, 2(1), 123-126.

Onarkulov, M., & Meliboyeva, A. (2024). HEMMING NEYRON TO'RLAR VA ULARNING ARXITEKTURASI. Current approaches and new research in modern sciences, 3(4), 177-181.

Onarkulov, M., & Satinova, G. (2024). NEYRON TO'RLARIDA FAOLLASHTIRISH FUNKSIYALARI. Theoretical aspects in the formation of pedagogical sciences, 3(8), 26-30.

Karimberdiyevich, O. M., & Mahamadamin o'g'li, Y. A. (2023). BASHORATLI TAHLILLAR UCHUN MASHINALI O'QITISH ALGORITMLARI. QIYOSIY QARASHLAR. Journal of Integrated Education and Research, 2(1), 130-134.

Foydalanilgan Veb-saytlar

1. MongoDB: <https://www.mongodb.com>
2. Cassandra: <https://cassandra.apache.org>
3. Neo4j: <https://neo4j.com>
4. Couchbase: <https://www.couchbase.com>