



MA'LUMOTLARNI SINFLASHTIRISHDA BIRCH ALGORITMI AHAMIYATI

Sanayev Mashrab Eshquvvat o'g'li

E-mail: sanayevmashrab@gmail.com

O'zbekiston - Finlandiya Pedagogika Instituti

Aniq fanlar kafedrası assistenti

Jalilov Shavkat Shodiyor o'g'li

E-mail: shavkatjalilov6465@gmail.com

O'zbekiston - Finlandiya Pedagogika Instituti

Toshtemirov Abdumalik Raxmonqul o'g'li

E-mail: abdumaliktoshtemirov@gmail.com

O'zbekiston Milliy Universiteti Jizzax filiali, Talaba

Annotatsiya

Bu maqola BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies) algoritmini ma'lumotlarni sinflashtirish uchun qanday qo'llanishni va uning ahamiyatini tahlil qiladi. Ma'lumotlar tizimlari kundalik hayotimizda o'rnatilgan vaqtincha ma'lumotlarning ko'plab qismini o'z ichiga oladi. Bu ko'plab ma'lumotlar bilan ishlashda, ularni tuzish va sinflashtirish muhimdir. BIRCH algoritmi, massiv ma'lumotlarini avlodlarga bo'lish va sinflashtirishda yaxshi natijalar olish uchun mo'ljallangan.

Kalit so'zlar: BIRCH, cluster feature, clusters, Cluster Feature Tree, Dendrogramma, Ballar soni, Chiziqli summalar, Kvadrat qiymatlar yig'indisi

KIRISH

BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies) algoritmi, massiv ma'lumotlarini tezkor va samarali ravishda sinflashtirish uchun mo'ljallangan algoritm bo'lib, avlodlarga bo'lish va ma'lumotlarni tuzishda yuqori samarali natijalar olishga xizmat qiladi. BIRCH algoritmi avlodlarni tuzishda "qavs" (cluster feature) va "qavslar" (clusters) deb nomlangan asosiy qo'llaniladigan elementlardan foydalanadi. Qavs, ma'lumotlar o'rtasidagi o'lcham va joylashuvi haqida umumiy ma'lumotlar to'plamini ifodalaydi. Avlodlar esa ma'lumotlar tuzilishini avtomatik ravishda ifodalaydi. Algoritm avlodlarni tuzishda "CF Tree" (Cluster Feature Tree) dan foydalanadi. Bu daraxt, boshlang'ich ma'lumotlarni RCC (root cluster feature) ga bog'lab, keyinchalik qavs va qavslarning qo'shimcha



ma'lumotlarini bog'lash orqali hosil qilinadi. Dendrogramma, avlodlarning bir-biri bilan o'zaro bog'lanishini va ma'lumotlar o'lchami va joylashuvi haqida ma'lumotlarni ko'rsatadi.

BIRCH algoritmi avlodlarni yaxshi taxlashda muhim ma'lumotlarni saqlaydi. Qavs va qavslarning o'lcham va joylashuvi xususiyatlari, avlodlarni ideal ravishda tuzishda yordam beradi. Algoritm dan olingan natijalar esa avlodlar va ma'lumotlar o'lchami bo'yicha o'zaro aloqani yaxshi aniqlashga imkon beradi. Shuningdek yirik ma'lumotlar bilan ishlashda samarali bo'lishi sababli, katta miqdordagi ma'lumotlar bilan ishlovchi. Bu esa masofaviy tashqaridagi ma'lumotlar, yozuvlar va avlodlar uchun idealdir. Algoritm ayni paytda tezkor va kichik xotira talabini bajaradi. BIRCH, ishtirokchi algoritmlar bilan hamqadam bo'lishi mumkin. Bu esa avlodlarni yaxshi tuzishda va sinflashtirishda yordam beradi. Algoritm, ma'lumotlarni tezkor va samarali ravishda sinflashtirishda muhim vazifalarni bajaradi. BIRCH algoritmi, massiv ma'lumotlarni sinflashtirishda avtomatik tuzish va yaxshi taxlash imkoniyatlarini ta'minlashda o'zini ko'rsatadi. U, yirik ma'lumotlar bilan ishlash, ishtirokchi algoritmlar bilan qo'llanish va kichik xotira talabini bajarish kabi asosiy qo'llaniladigan elementlarga ega bo'lib, masofaviy tashqaridagi ma'lumotlarni boshqarishda samarali yechimdir.

BIRCH algoritmidagi CF ning ahamiyati nimada?

BIRCH zich hududlarda joylashgan ma'lumotlarni Klasterlash xususiyati (CF) yozuvlari sifatida umumlashtirish orqali katta ma'lumotlar to'plamining xotira talablarini minimallashtirishga harakat qiladi. CF daraxti ma'lumotlar to'plamining juda ixcham tasviridir, chunki barg tugunidagi har bir yozuv bitta ma'lumot nuqtasi emas, balki kichik klasterdir. Har bir barg bo'lmagan tugunda eng ko'p B yozuvlari mavjud. Shu nuqtai nazardan, bitta yozuvda bola tuguniga ko'rsatgich va boladagi CF lar yig'indisidan tashkil topgan CF (kichik klasterlarning kichik guruhleri) mavjud. Boshqa tomondan, barg tugunida eng ko'p L yozuvlar mavjud va har bir yozuv CF (ma'lumotlar nuqtalarining kichik guruhleri) hisoblanadi. Barg tugunidagi barcha yozuvlar chegara talabini qondirishi kerak. Ya'ni, har bir barg kirishining diametri Eshik qiymatidan kichik bo'lishi kerak. Bundan tashqari, har bir barg tugunida oldingi va keyingi ikkita ko'rsatkich mavjud bo'lib, ular samarali skanerlash uchun barcha barg tugunlarini bir-biriga bog'lash uchun ishlatiladi.

BIRCH klasterlashda qanday statistik ma'lumotlardan foydalaniladi?

BIRCH tomonidan qo'llaniladigan klasterlash xususiyatlari - bu yangi ma'lumotlar bilan osongina yangilanishi mumkin bo'lgan oddiy umumiy statistika:

Ballar soni,



Chiziqli summalar,
Kvadrat qiymatlar yig'indisi.

FOYDALANILGAN ADABIYOTLAR:

1. Naim o'g'li M. D., Abdishukur o'g'li S. A. THE NUMPY LIBRARY OF THE PYTHON PROGRAMMING LANGUAGE IS AN OPTIMAL SOLUTION FOR WORKING WITH ARRAYS //JOURNAL OF INNOVATIONS IN SCIENTIFIC AND EDUCATIONAL RESEARCH. – 2022. – Т. 2. – №. 13. – С. 195-197.
2. Naim o'g'li M. D., Baxtiyor o'g'li E. S. DATA SCIENCE METHODOLOGY IN LEARNING PROGRAMMING //JOURNAL OF INNOVATIONS IN SCIENTIFIC AND EDUCATIONAL RESEARCH. – 2022. – Т. 2. – №. 13. – С. 207-210.
3. Amanbayevna A. S., Naim o'g'li M. D. GEOMETRIC MODELING AND VISUALIZATION OF SELF-SIMILAR STRUCTURES BASED ON FRACTAL THEORY //JOURNAL OF INNOVATIONS IN SCIENTIFIC AND EDUCATIONAL RESEARCH. – 2022. – Т. 2. – №. 13. – С. 187-188.
4. Naim o'g'li M. D., Baxtiyor o'g'li E. S. KATTA HAJMDAGI MA'LUMOTLARINI TAHLIL QILISHDA CRISP-DM USULLARINI QO'LLASH //JOURNAL OF INNOVATIONS IN SCIENTIFIC AND EDUCATIONAL RESEARCH. – 2022. – Т. 2. – №. 13. – С. 211-216.
5. Baxtiyor o'g'li E. S., Naim o'g'li M. D. YAGONA INTERAKTIV DAVLAT XIZMATLARI PORTALNING TRANSPORT BO'LIMINI IDF0, DFD VA IDF3 SATANDART DIAGRAMMARI YORDAMIDA YARATILGAN LOYIHASI //JOURNAL OF INNOVATIONS IN SCIENTIFIC AND EDUCATIONAL RESEARCH. – 2022. – Т. 2. – №. 13. – С. 198-206.
6. Doston M., Abdulatif S. SUN'IY INTELLEKT TEXNOLOGIYALARI VA ULARNI SOHALARDA QO'LLANILISHI //Educational Research in Universal Sciences. – 2022. – Т. 1. – №. 5. – С. 144-147.
7. Javlon X. et al. Классификатор движения рук с использованием биомиметического распознавания образов с помощью сверточных нейронных сетей с методом динамического порога для извлечения движения с использованием датчиков EF //Journal of new century innovations. – 2022. – Т. 19. – №. 6. – С. 352-357.
8. Naim o'g'li M. D., Xasan o'g'li A. M. CLASSIFICATION OF CRISP-DM METHODOLOGY IN DATA ANALYSIS //MODELS AND METHODS FOR INCREASING THE EFFICIENCY OF INNOVATIVE RESEARCH. – 2023. – Т. 2. – №. 19. – С. 220-222.
9. Naim o'g'li, Muxtorov Doston, and Qudratov Akbar Akmal o'g'li. "ROLAP VA DATA MINING INTEGRATSIYASI." *Yangi O'zbekiston taraqqiyotida tadqiqotlarni o'rni va rivojlanish omillari* 1.1 (2023): 233-236.



10. Naim o'g'li, Muxtorov Doston, and RASHIDOV BOYSARI TO'RA O'G. "STREAM DATA MINING VA REAL VAQT TAHLILI." *Yangi O'zbekiston taraqqiyotida tadqiqotlarni o'rni va rivojlanish omillari* 1.1 (2023): 217-222.
11. Naim o'g'li, Muxtorov Doston, ALISHEROV BOBURJON ANVAR O'G'LI, and SAFAROV ULUG'BEK ANORBOY O'G. "OLAP VA DATA MINING INTEGRATSIYASI." *Yangi O'zbekiston taraqqiyotida tadqiqotlarni o'rni va rivojlanish omillari* 1.1 (2023): 229-232.
12. Naim o'g'li, Muxtorov Doston, SAFAROV ULUG'BEK ANORBOY O'G, and G'ANIYEV SOBIR QODIR O'G. "DATA MINING KLASSIFIKATSIYA VA KLASTERIZATSIYA." *Yangi O'zbekiston taraqqiyotida tadqiqotlarni o'rni va rivojlanish omillari* 1.1 (2023): 223-228.
13. Naim o'g'li, Muxtorov Doston, and Shermuhammedov Abdulatif Abdishukur o'g'li. "THE NUMPY LIBRARY OF THE PYTHON PROGRAMMING LANGUAGE IS AN OPTIMAL SOLUTION FOR WORKING WITH ARRAYS." *JOURNAL OF INNOVATIONS IN SCIENTIFIC AND EDUCATIONAL RESEARCH* 5.4 (2022): 195-197.
14. Naim o'g'li, Muxtorov Doston, and Ergashev Sirojiddin Baxtiyor o'g'li. "KATTA HAJMDAGI MA'LUMOTLARINI TAHLIL QILISHDA CRISP-DM USULLARINI QO'LLASH." *JOURNAL OF INNOVATIONS IN SCIENTIFIC AND EDUCATIONAL RESEARCH* 5.4 (2022): 211-216.
15. Baxtiyor o'g'li, Ergashev Sirojiddin, and Muxtorov Doston Naim o'g'li. "YAGONA INTERAKTIV DAVLAT XIZMATLARI PORTALNING TRANSPORT BO'LIMINI IDF0, DFD VA IDF3 SATANDART DIAGRAMMARI YORDAMIDA YARATILGAN LOYIHASI." *JOURNAL OF INNOVATIONS IN SCIENTIFIC AND EDUCATIONAL RESEARCH* 5.4 (2022): 198-206.
16. Yusupovich X. J. BEMORLARNING SHIFOKOR YOZGAN RETSEPTI BOYICHA DORILARNI QABUL QILGANLIK DARAJASINI ANIQLASH AVTOMATIK TIZIMNI YARATISH //MODELS AND METHODS FOR INCREASING THE EFFICIENCY OF INNOVATIVE RESEARCH. – 2023. – Т. 2. – №. 19. – С. 223-234.
17. Бурнашев В. Ф., Холматов Ж. Ю. МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ГИДРОДИНАМИЧЕСКИХ ПРОЦЕССОВ МНОГОФАЗНОЙ ФИЛЬТРАЦИИ В НЕФТЯНОМ ПЛАСТЕ ПРИ ЕГО ЗАВОДНЕНИИ //RESEARCH AND EDUCATION. – 2022. – Т. 1. – №. 1. – С. 137-154.
18. Kholmatov Javlon, & Mustafoyev Erali. (2023). STRUCTURE AND PRINCIPLE OF OPERATION OF FULLY CONNECTED NEURAL NETWORKS. *International Journal of Contemporary Scientific and Technical Research*, 136–141.