

AJRATUVCHI GIPER TEKISLIKLAR VA ULARNING MASHINALI O'QITISH ALGORITMLARIDA QO'LLANILISHI

Tojimamatov Israil Nurmamatovich

Fargʻona davlat universiteti

Amaliy matematika va informatika

kafedrası katta o'qituvchisi

E-mail: israiltojimatov@gmail.com

Omonjonova Mavludaxon Kamoliddin qizi

Fargʻona davlat universiteti Amaliy

matematika yoʻnalishi 3-bosqich 23.08-

guruh talabasi

E-mail: ozodbekomonjonov203@gmail.com

ANNOTATSIYA: Ushbu tadqiqot ishida ajratuvchi giper tekisliklarning nazariy asoslari hamda ularning zamonaviy mashinali o'qitish algoritmlarida qo'llanilishi chuqur tahlil qilingan. Giper tekisliklar ko'p o'lchovli fazolarda ikki yoki undan ortiq sinfni ajratishda muhim rol o'ynaydi va turli xil klassifikatsiya muammolarini hal qilishda keng qo'llaniladi. Maqolada giper tekisliklarning geometrik xususiyatlari, ularning optimal tanlanish mezonlari hamda amaliy dasturlarda qo'llanilishi masalalari o'rganilgan. Maxsus e'tibor qo'llab-vektor mashinalari usulida qo'llaniladigan maksimal marginal giper tekisliklar konsepsiyasiga qaratilgan. Tadqiqot doirasida chiziqli ajraladigan va chiziqli ajralmadigan to'plamlar uchun turli yondashuvlar tahlil qilinib, kernel funksiyalaridan foydalanish orqali murakkab muammolarni hal qilish imkoniyatlari ko'rsatilgan. Olingan natijalar turli predmetli sohalarda klassifikatsiya muammolarini samarali hal qilish uchun nazariy va amaliy asos yaratadi.

Kalit so'zlar: ajratuvchi giper tekislik, mashinali o'qitish, klassifikatsiya, qo'llabvektor mashinalari, maksimal marginal, kernel usullari, chiziqli ajratish, ko'p o'lchovli fazo

ANNOTATION: This research paper provides an in-depth analysis of the theoretical foundations of separating hyperplanes and their applications in modern machine learning algorithms. Hyperplanes play a crucial role in separating two or more classes in multidimensional spaces and are widely applied in solving various classification problems. The article examines the geometric properties of hyperplanes, their optimal selection criteria, and issues related to their application in practical settings. Special attention is given to the concept of maximum margin hyperplanes used in support vector machine methods. Within the scope of the research, different approaches for linearly separable and non-linearly separable datasets have been analyzed, and the possibilities of solving complex problems through the use of kernel functions have been demonstrated. The obtained results create a theoretical and practical foundation for effectively solving classification problems in various subject domains.

Keywords: separating hyperplane, machine learning, classification, support vector machines, maximum margin, kernel methods, linear separation, multidimensional space

АННОТАЦИЯ: В данной исследовательской работе проведен углубленный анализ теоретических основ разделяющих гиперплоскостей и их применения в современных алгоритмах машинного обучения. Гиперплоскости играют важнейшую роль в разделении двух или более классов в многомерных пространствах и широко применяются при решении различных задач классификации. В статье рассмотрены геометрические свойства гиперплоскостей, критерии их оптимального выбора и вопросы применения в практических приложениях. Особое внимание уделено концепции гиперплоскостей с максимальным зазором, используемых в методе опорных векторов. В рамках исследования проанализированы различные подходы для линейно разделимых и линейно неразделимых множеств, продемонстрированы

возможности решения сложных задач посредством использования ядерных функций. Полученные результаты создают теоретическую и практическую основу для эффективного решения задач классификации в различных предметных областях.

Ключевые слова: разделяющая гиперплоскость, машинное обучение, классификация, метод опорных векторов, максимальный зазор, ядерные методы, линейное разделение, многомерное пространство

KIRISH

Hozirgi zamon axborot texnologiyalari rivojlanishining jadal sur'atlari ma'lumotlarni tahlil qilish va klassifikatsiya qilish muammolarini dolzarbligini oshirmoqda. Katta hajmdagi ma'lumotlar oqimini qayta ishlash zarurati turli sohalarda samarali klassifikatsiya algoritmlarini yaratishni talab etadi. Mashinali o'qitish sohasida ajratuvchi giper tekisliklar konsepsiyasi asos tosh bo'lib xizmat qiladi va ko'plab zamonaviy algoritmlarning nazariy bazasini tashkil etadi.

Giper tekislik tushunchasi ko'p o'lchovli geometriyadan kelib chiqadi va n -o'lchovli fazoda n -bir o'lchovli affin ostfazoni ifodalaydi. Oddiy qilib aytganda, ikki o'lchovli fazoda giper tekislik to'g'ri chiziqli, uch o'lchovli fazoda esa tekislik bo'ladi. Ko'p o'lchovli fazolarda esa giper tekislik murakkab geometrik ob'ekt sifatida namoyon bo'ladi va turli sinflarga tegishli nuqtalarni ajratish imkonini beradi.

Klassifikatsiya muammolari zamonaviy sun'iy intellekt tizimlarining markaziy yo'nalishlaridan biridir. Tibbiy diagnostika, moliyaviy tahlil, tasvirlarni tanish, tabiiy tilni qayta ishlash kabi sohalarda ma'lumotlarni to'g'ri sinfga ajratish muammosi doimiy ravishda yuzaga keladi. Masalan, tibbiy tasvirlarni tahlil qilishda kasallik belgilarini aniqlash, elektron pochta xabarlarini spam va oddiy xabarlariga ajratish, yoki kredit berish tizimlarida mijozlarning ishonchliligini baholash jarayonlarida klassifikatsiya algoritmlari muhim ahamiyat kasb etadi.

Ajratuvchi giper tekisliklar nazariyasi asosan ikki asosiy muammoni hal qilishga qaratilgan. Birinchidan, berilgan ma'lumotlar to'plamini ikki yoki undan ortiq sinfga ajratuvchi giper tekislikning mavjudligini aniqlash zarur. Ikkinchidan, agar bunday giper

tekislik mavjud bo'lsa, uni optimal tanlash mezonlarini ishlab chiqish va amalda qo'llash usullarini yaratish kerak. Bu muammolar geometrik, algebraik va optimallashtirish nazariyalarining uyg'unligini talab etadi.

Zamonaviy tadqiqotlar giper tekisliklarning turli xususiyatlarini chuqur o'rganishga qaratilgan. Maksimal marginal konsepsiyasi qo'llab-vektor mashinalarining asosiy g'oyasi bo'lib, u ajratuvchi giper tekislikni shunday tanlashni nazarda tutadiki, turli sinflarga tegishli eng yaqin nuqtalar orasidagi masofa maksimal bo'lsin. Bu yondashuv modelning umumlashtirish qobiliyatini oshiradi va yangi ma'lumotlarni klassifikatsiya qilishda yuqori aniqlikni ta'minlaydi.

Ushbu tadqiqot ishida yuqoridagi savollarga javob topish va ajratuvchi giper tekisliklar nazariyasini keng qamrovli tahlil qilish maqsad qilib qo'yilgan. Ishda giper tekisliklarning geometrik va algebraik xususiyatlari batafsil o'rganiladi, optimal ajratish mezonlari tahlil qilinadi va zamonaviy mashinali o'qitish algoritmlarida ularning qo'llanilishi ko'rsatiladi. Maxsus e'tibor qo'llab-vektor mashinalari va kernel usullariga qaratiladi, chunki bu usullar hozirgi kunda eng samarali klassifikatsiya vositalaridan **hisoblanadi.**

NAZARIY ASOS: GIPER TEKISLIKLAR MATEMATIKASI

Giper tekisliklar nazariyasini chuqur tushunish uchun avvalo ularning matematik ifodasini va xususiyatlarini batafsil ko'rib chiqish zarur. n -o'lchovli Evklid fazosida giper tekislik quyidagi chiziqli tenglama bilan ifodalanadi:

$$w_1x_1 + w_2x_2 + \dots + w_nx_n + b = 0$$

Bu yerda $w = (w_1, w_2, \dots, w_n)$ vektor giper tekislikka perpendikulyar bo'lgan normal vektor hisoblanadi, $x = (x_1, x_2, \dots, x_n)$ fazodagi ixtiyoriy nuqta koordinatalari, b esa qo'shimcha o'zgarma bo'lib, giper tekislikning koordinatalar boshiga nisbatan joylashuvini belgilaydi. Normal vektor w giper tekislikning yo'nalishini to'liq aniqlaydi va parametr b esa uni fazoda siljitish imkonini beradi.

QO'LLAB-VEKTOR MASHINALARI NAZARIYASI

Qo'llab-vektor mashinalari ajratuvchi giper tekisliklar nazariyasining eng muvaffaqiyatli amaliy tatbiqi hisoblanadi. Ushbu usul statistik o'qitish nazariyasiga asoslanadi va maksimal marginal printsipini qo'llaydi. Qo'llab-vektor mashinalari klassifikatsiya muammolarida yuqori aniqlik va yaxshi umumlashtirish qobiliyati bilan ajralib turadi.

Qo'llab-vektor mashinalari algoritmining asosiy g'oyasi shundan iboratki, o'quv to'plamini ajratuvchi barcha mumkin bo'lgan giper tekisliklar ichidan marginal eng katta bo'lganini tanlash kerak. Bu yondashuv nazariy asosga ega va Vapnik-Chervonenkis nazariyasi doirasida qat'iy isbotlangan. Maksimal marginal umumlashtirish xatosining yuqori chegarasini kamaytiradi va modelning yangi Nazariy tahlillar shuni ko'rsatadiki, maksimal marginalga ega ajratuvchi giper tekislik generalizatsiya xatosining eng kichik yuqori chegarasini beradi. Margin qancha katta bo'lsa, model yangi ma'lumotlarga shuncha yaxshi umumlashadi. Bu xususiyat qo'llabvektor mashinalari usulini boshqa klassifikatsiya usullaridan ajratib turadi va uning muvaffaqiyatini tushuntiradi.

KERNEL USULLARI VA CHIZIQLI AJRALMAGAN TO'PLAMLAR

Ko'plab amaliy muammolarda ma'lumotlar chiziqli ajraladigan bo'lmaydi. Bunday hollarda ma'lumotlarni yuqori o'lchovli fazoga o'tkazish va u yerda chiziqli ajratish mumkin bo'ladi. Kernel usullari bu g'oyani samarali amalga oshirish imkonini beradi va murakkab klassifikatsiya muammolarini hal qilishda keng qo'llaniladi.

Kernel parametrlarini tanlash modelning samaradorligiga katta ta'sir ko'rsatadi. Gauss kerneli uchun γ parametri modelning murakkabligini boshqaradi. Kichik γ qiymati yumshoq qaror chegaralariga olib keladi va modelning ortiqcha o'rganish xavfini kamaytiradi. Katta γ qiymati esa murakkab qaror chegaralari yaratadi va o'quv to'plamiga yaqinroq moslanadi, lekin yangi ma'lumotlarga yomon umumlashishi mumkin.

Polinom kerneli uchun d daraja parametri muhim rol o'ynaydi. Kichik darajalar oddiy munosabatlarni, katta darajalar esa murakkab o'zaro ta'sirlarni ifodalaydi. Amalda odatda

ikkinchi yoki uchinchi darajali polinom kernellari qo'llaniladi, chunki yuqori darajali polinomlar hisoblash jihatidan qimmat va ortiqcha o'rganish xavfini oshiradi.

Kernel usullari chiziqli ajralmagan ma'lumotlar bilan ishlashda juda samarali bo'lsada, bir qancha cheklovlarga ega. Birinchidan, kernel matritsasini hisoblash va saqlash katta hajmli to'plamlar uchun xotira talab qiladi. Kernel matritsasining o'lchami $n \times n$ bo'lib, bu yerda n o'quv namunalari soni. Millionlab nuqtadan iborat to'plamlar uchun bu jiddiy muammo bo'lishi mumkin.

Ikkinchidan, optimal kernel turini va uning parametrlarini tanlash qiyin vazifa hisoblanadi. Turli kernellarni sinab ko'rish va kross-validatsiya orqali baholash vaqt talab etadi. Uchinchidan, kernel usullari bilan yaratilgan modellar talqin qilish jihatidan qiyinroq bo'ladi, chunki qaror chegarasi yuqori o'lchovli fazoda ifodalanadi.

Kernel usullarining yana bir muhim jihati shundaki, ular faqat qo'llab-vektor mashinalari bilan cheklanmaydi. Boshqa mashinali o'qitish algoritmlariga ham qo'llanilishi mumkin. Masalan, kernel PCA usuli o'lchovlarni kamaytirish uchun, kernel regression usuli regressiya muammolarida qo'llaniladi. Kernel usullari umuman olganda chiziqli usullarni nochiziqli muammolarga kengaytirish imkonini beruvchi kuchli vosita hisoblanadi.

AMALIY QO'LLANMALAR

Ajratuvchi giper tekisliklar va qo'llab-vektor mashinalari zamonaviy sun'iy intellekt tizimlarida keng qo'llaniladi. Ularning samaradorligi turli sohalarda isbotlangan va ko'plab muvaffaqiyatli loyihalar amalga oshirilgan.

Tibbiy diagnostika sohasida qo'llab-vektor mashinalari kasalliklarni aniqlash va prognozlashda qo'llaniladi. Tibbiy tasvirlarni tahlil qilishda, masalan, rentgen, MRI yoki KT tasvirlarida patologik o'zgarishlarni aniqlashda yuqori aniqlikka erishilgan. Rak kasalligini erta bosqichda aniqlash uchun qo'llab-vektor mashinalari asosidagi tizimlar ishlab chiqilgan. Bemorlarning klinik ma'lumotlari asosida kasallik xavfini baholash va davolash strategiyasini tanlashda ham ushbu usullar samarali ekanligini ko'rsatgan.

Molekulyar biologiya va bioinformatika sohalarida protein strukturasini bashorat qilish, genlarni klassifikatsiya qilish va oqsil-oqsil o'zaro ta'sirini aniqlashda qo'llab-vektor

mashinalari qo'llaniladi. Genom ketma-ketliklarini tahlil qilishda va mutatsiyalarning ta'sirini baholashda ham ushbu usullar keng tatbiq etiladi. DNK mikroarray ma'lumotlarini qayta ishlashda qo'llab-vektor mashinalari yuqori o'lchovli xususiyatlar bilan samarali ishlash qobiliyatini namoyish etadi.

Moliyaviy tahlil sohasida kredit risklarini baholash, qimmatli qog'ozlar bozorini bashorat qilish va firibgarlikni aniqlashda qo'llab-vektor mashinalari qo'llaniladi. Bank tizimlarida mijozlarning kredit layoqatligini baholash uchun qo'llab-vektor mashinalari asosidagi modellar ishlab chiqilgan. Bu modellar tarixiy ma'lumotlardan o'rganib, yangi mijozlar uchun xavf darajasini aniq baholaydi. To'lov kartalaridan foydalanishda firibgarlik operatsiyalarini real vaqt rejimida aniqlash tizimlari ham ushbu texnologiyaga asoslanadi.

Tasvirlarni tanish va kompyuter ko'rishi sohalarida qo'llab-vektor mashinalari ob'ektlarni klassifikatsiya qilish, yuzlarni tanish va belgilarni ajratishda qo'llaniladi. Avtomobil raqamlarini avtomatik tanish tizimlari, xavfsizlik kameralarida odamlarni identifikatsiyalash va atrof-muhitni tahlil qilish tizimlarida ushbu usullar asosiy rol o'ynaydi. Tibbiy tasvirlarni qayta ishlashda, masalan, rentgen tasvirlarida o'pka kasalliklarini yoki suyak siniqlarini aniqlashda yuqori aniqlikka erishilgan.

Xavfsizlik tizimlarida kirish nazorati, hodisalarni aniqlash va xavflarni baholashda qo'llab-vektor mashinalari qo'llaniladi. Kiberhavfsizlik sohasida tarmoq trafigini tahlil qilib, hujumlarni aniqlash tizimlari yaratilgan. Bu tizimlar anomal harakatlarni real vaqt rejimida aniqlab, zararli faoliyatni to'xtatadi.

ALGORITMNING HISOBLASH MURAKKABLIGI VA OPTIMALLASHTIRISH

Qo'llab-vektor mashinalari algoritmini amalda qo'llashda hisoblash murakkabligi muhim masala hisoblanadi. Katta hajmli ma'lumotlar to'plamlari bilan ishlashda samarali algoritmlar va optimallashtirish usullarini qo'llash zarur.

Klassik kvadratik optimallashtirish usullari bilan dual masalani hal qilishning hisoblash murakkabligi o'quv namunalari soni n ga nisbatan $O(n^3)$ tartibida bo'ladi. Bu

degani, ma'lumotlar hajmi ikki marta oshsa, hisoblash vaqti sakkiz marta ortadi. Minglab yoki millionlab namunalardan iborat to'plamlar uchun bunday yondashuv amaliy emas.

Sequential Minimal Optimization algoritmi hisoblash murakkabligini sezilarli darajada kamaytiradi. Bu algoritim har safar faqat ikkita Lagranj ko'paytuvchisini optimallashtirib, qolganlarini o'zgarishsiz qoldiradi. Ikki o'zgaruvchili masalani analitik hal qilish mumkin va bu jarayon tez amalga oshiriladi. Algoritmnining umumiy murakkabligi taxminan $O(n^2)$ tartibida bo'lib, amalda yanada yaxshi ishlaydi.

Kernel matritsasini hisoblash va saqlash ham katta resurs talab qiladi. $n \times n$ o'lchamdagi kernel matritsasi n^2 ta elementga ega va har biri kernel funksiyasini hisoblashni talab etadi. Katta n qiymatlari uchun bu jiddiy muammo bo'lishi mumkin. Yechim sifatida kesh usullaridan foydalaniladi: tez-tez ishlatiladigan kernel qiymatlari xotirada saqlanadi, qolganlari esa kerak bo'lganda hisoblanadi.

Parallel hisoblash usullari qo'llab-vektor mashinalari algoritmini tezlashtirish uchun qo'llaniladi. Kernel matritsasining elementlarini parallel hisoblash mumkin, chunki ular bir-biriga bog'liq emas. Zamonaviy ko'p yadroli protsessorlar va GPU lar yordamida hisoblash vaqtini sezilarli darajada qisqartirish mumkin.

Approksimatsiya usullari ham qo'llaniladi. Kernel funksiyasini past darajali matritsa bilan yaqinlashtirib, hisoblash murakkabligini kamaytirish mumkin. Random Features usuli va Nystrom usuli keng tarqalgan approksimatsiya yondashuvlaridir. Bu usullar aniqlikni ozgina kamaytirish evaziga hisoblash vaqtini sezilarli qisqartiradi.

Onlayn o'qitish algoritmlari oqim rejimida kelayotgan ma'lumotlar bilan ishlash imkonini beradi. Bu yondashuvda model yangi ma'lumotlar kelgani sari yangilanadi va butun tarixiy ma'lumotlarni xotirada saqlash zarurati yo'qoladi. Stochastic gradient descent usuli va uning turli xil modifikatsiyalari onlayn o'qitish uchun qo'llaniladi.

Xususiyatlarni tanlash va o'lchovlarni kamaytirish ham hisoblash samaradorligini oshiradi. Ortiqcha yoki ahamiyatsiz xususiyatlarni olib tashlash kernel matritsasining o'lchamini kamaytiradi va algoritmni tezlashtiradi. Principal Component Analysis yoki boshqa o'lchovlarni kamaytirish usullaridan foydalanish mumkin.

XULOSA

Ushbu tadqiqot ishida ajratuvchi giper tekisliklarning nazariy asoslari va zamonaviy mashinali o'qitish algoritmlarida qo'llanilishi keng qamrovli tahlil qilindi. Tadqiqot natijalari shuni ko'rsatdiki, giper tekisliklar ko'p o'lchovli fazolarda klassifikatsiya muammolarini hal qilishda fundamental rol o'ynaydi va qo'llab-vektor mashinalari usuli ularning eng samarali tatbiqlaridan biri hisoblanadi.

Giper tekisliklarning geometrik va algebraik xususiyatlari to'liq o'rganildi. No'lchovli Evklid fazosida giper tekislik chiziqli tenglama bilan ifodalanishi va fazoni ikki yarim fazoga ajratishi aniqlandi. Normal vektor va siljish parametri giper tekislikning pozitsiyasi va yo'nalishini to'liq belgilashi ko'rsatildi. Nuqtadan giper tekislikkacha bo'lgan masofani hisoblash formulasi va uning klassifikatsiya muammolarida qo'llanilishi batafsil tahlil qilindi.

Kelajakda tadqiqotlar bir necha yo'nalishda davom etishi kutilmoqda. Birinchidan, hisoblash samaradorligini yanada oshirish va juda katta hajmli ma'lumotlar bilan ishlash imkoniyatlarini kengaytirish. Ikkinchidan, murakkab strukturali ma'lumotlar bilan ishlash qobiliyatini rivojlantirish. Uchinchidan, modellarning izohlanuvchanligi va shaffofligini oshirish. To'rtinchidan, noxolis qarorlar muammosini hal qilish va adolatli mashinali o'qitish usullarini yaratish.

Ajratuvchi giper tekisliklar nazariyasi va qo'llab-vektor mashinalari zamonaviy mashinali o'qitishning fundamental asoslaridan biri bo'lib, yangi algoritmlar va usullarni yaratishda ilhom manbai bo'lib qolmoqda. Ushbu tadqiqot ishi ajratuvchi giper tekisliklarning nazariy va amaliy jihatlarini keng qamrovli tahlil qilish orqali bu sohani o'rganadiganlar uchun to'liq ma'lumot bazasini yaratdi va kelajakdagi tadqiqotlar uchun asos yaratdi.

FOYDALANILGAN ADABIYOTLAR

1. Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers. Proceedings of the fifth annual workshop on Computational learning theory. 1992; 144152.
2. Cortes C, Vapnik V. Support-vector networks. Machine Learning. 1995; 20(3): 273297.

3. Vapnik VN. The Nature of Statistical Learning Theory. Springer-Verlag New York. 1995.
4. Cristianini N, Shawe-Taylor J. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press. 2000.
5. Scholkopf B, Smola AJ. Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. MIT Press. 2002.
6. Tojimamatov, I. N., & Saidjamolova, B. M. (2023). BIZNESDA «BIG DATA» TEXNOLOGIYALARI VA ULARNING AHAMIYATI. Лучшие интеллектуальные исследования, 11(4), 56–63.
7. Tojimamatov, I. N., & Azizjon o'g'li, N. A. Z. (2024). The SQL server language and its structure. American Journal of Open University Education, 1(1), 11–15.
8. Tojimamatov, I. N., Topvoldiyeva, H., Karimova, N., & Inomova, G. (2023). GRAFIK MA'LUMOTLAR BAZASI. Евразийский журнал технологий и инноваций, 1(4), 75–84.
9. Tojimamatov, I. N., & Gulhayo, M. (2023). MA'LUMOTLARNI QAYTA ISHLASHDA ERP TIZIMLARI. MA'LUMOTLARNI QAYTA ISHLASHDA SAP TIZIMLARI. Journal of Integrated Education and Research, 2(4), 87–89.