

ALTSGEYMER KASALLIGINI ERTA ANIQLASHDA MASHINALI O'RGATISH ALGORITMLARINI QO'LLANILISHI

Abdunabiyeva Diyora Abdug`affor qizi

Mirzo Ulug`bek nomidagi O'zbekiston Milliy Universtieti

Sun'iy intellekt mutaxassisligi magistranti

+998947312535

abdunabiyevadiyora048@gmail.com

Abstract. This study proposes a multi-criteria evaluation approach for assessing machine learning models in the early detection of Alzheimer's disease. Conventional evaluation based on individual metrics such as accuracy, sensitivity, specificity, and ROC-AUC does not fully reflect overall diagnostic performance. To address this issue, an integral evaluation metric is introduced by combining multiple performance measures into a single indicator. The approach is applied to several models, including Random Forest, Support Vector Machine, and Convolutional Neural Network, using clinical and imaging data. The results show that the proposed method provides a more balanced and reliable assessment of model performance, enabling effective model selection. The approach can be utilized in medical decision-support systems to improve the accuracy and reliability of early diagnosis.

Keywords: alzheimer's disease, machine learning, model evaluation, integrated metric, multi-criteria analysis

Аннотация. В данной работе предлагается многокритериальный подход к оценке эффективности моделей машинного обучения для раннего выявления болезни Альцгеймера. Традиционные методы оценки, основанные на отдельных метриках, таких как точность, чувствительность, специфичность и ROC-AUC, не позволяют в полной мере отразить общую диагностическую способность модели. Для решения данной проблемы предложен интегральный показатель, объединяющий несколько метрик в единый критерий оценки. Подход применён к различным моделям, включая Random Forest, Support Vector Machine и Convolutional Neural Network, с использованием клинических и визуальных данных. Полученные результаты показывают, что предложенный метод обеспечивает более сбалансированную и надёжную оценку моделей, что позволяет повысить обоснованность выбора оптимальной модели. Данный подход может быть эффективно использован в системах поддержки принятия медицинских решений.

Ключевые слова: болезнь Альцгеймера, машинное обучение, оценка моделей, интегральный показатель, многокритериальный анализ.

Annotatsiya. Mazkur ishda Altsgeymer kasalligini erta aniqlashda mashinali o'rganish modellarining samaradorligini baholash uchun ko'p mezonli yondashuv taklif etiladi. An'anaviy baholash usullari, xususan aniqlik, sezgirlik, spetsifiklik va ROC-AUC kabi alohida metrikalarga asoslangan yondashuvlar modelning umumiy diagnostik imkoniyatini to'liq aks ettirmaydi. Ushbu muammoni hal etish maqsadida bir nechta baholash mezonlarini yagona integral ko'rsatkichga birlashtiruvchi metod taklif etildi. Yondashuv Random Forest, Support Vector Machine va Convolutional Neural Network modellariga klinik va tasviriy ma'lumotlar asosida qo'llanildi.

Natijalar shuni ko'rsatdiki, taklif etilgan metod modellar samaradorligini yanada muvozanatli va ishonchli baholash imkonini beradi hamda optimal modelni tanlash jarayonini yaxshilaydi. Ushbu yondashuv tibbiy qaror qabul qilish tizimlarida samarali qo'llanishi mumkin.

Kalit so'zlar: altsgeymer kasalligi, mashinali o'rganish, modelni baholash, integral ko'rsatkich, ko'p mezonli tahlil

Kirish

Altsgeymer kasalligi (AK) markaziy asab tizimining degenerativ kasalliklari orasida eng keng tarqalgan demensiya shakli bo'lib, u kognitiv funksiyalar, xotira va fikrlash qobiliyatining bosqichma-bosqich susayishi bilan tavsiflanadi. Jahon sog'liqni saqlash tashkiloti hisobotlariga ko'ra, demensiya bilan kasallangan bemorlar soni global miqyosda ortib borayotganligi sababli, kasallikni erta aniqlash va samarali diagnostika usullarini ishlab chiqish dolzarb ilmiy muammo hisoblanadi [1].

An'anaviy diagnostika yondashuvlari, jumladan neyropsixologik testlar, magnit-rezonans tomografiya (MRT) va pozitron-emission tomografiya (PET) usullari, ko'pincha kasallikning nisbatan kech bosqichlarida yuqori aniqlik beradi [2]. Biroq, kasallikning boshlang'ich bosqichlarida yuzaga keladigan nozik o'zgarishlarni aniqlash murakkab bo'lib qolmoqda. Shu sababli so'nggi yillarda sun'iy intellekt va mashinali o'rganish usullariga asoslangan yondashuvlar tibbiy diagnostika tizimlarida keng qo'llanila boshladi [3].

Xususan, Machine Learning algoritmlari katta hajmdagi klinik, tasviriy va genetik ma'lumotlarni kompleks tahlil qilish orqali yashirin naqshlarni aniqlash imkonini beradi. Tadqiqotlar shuni ko'rsatadiki, Random Forest, Support Vector Machine va Convolutional Neural Network kabi modellar Altsgeymer kasalligini aniqlashda yuqori samaradorlikka ega [4–6]. Ayniqsa, MRT tasvirlari asosida chuqur o'rganish modellari yordamida kasallikni erta bosqichda aniqlash imkoniyatlari kengaymoqda [5].

Shu sababli mashinali o'rganish modellarini baholashda bir nechta mezonlarni kompleks tarzda hisobga oluvchi integrallashgan yondashuvlarni ishlab chiqish zarurati mavjud. Mazkur tadqiqotning asosiy maqsadi Altsgeymer kasalligini erta aniqlashda qo'llaniladigan mashinali o'rganish modellarining samaradorligini

baholash uchun bir nechta diagnostik metrikalarni yagona integral ko'rsatkichga birlashtiruvchi yondashuvni taklif etishdan iborat. Ushbu yondashuv turli algoritmlarni ob'ektiv taqqoslash imkonini berib, eng muvozanatli va klinik jihatdan ishonchli modelni aniqlashga xizmat qiladi.

1. Tadqiqot metodlari

Mazkur tadqiqotning asosiy maqsadi Altsgeymer kasalligini erta aniqlashda qo'llaniladigan mashinali o'rganish modellarining samaradorligini baholash jarayonini yanada ob'ektiv va kompleks shaklga keltirishdan iborat. So'nggi yillarda sun'iy intellektga asoslangan yondashuvlar tibbiy diagnostika tizimlarida keng qo'llanilib, kasalliklarni erta bosqichda aniqlashda yuqori natijalar ko'rsatmoqda [1,2].

Biroq mavjud tadqiqotlarda modellar samaradorligi odatda alohida baholash metrikalari — aniqlik (accuracy), sezgirlik (sensitivity), spetsifiklik (specificity) va ROC-AUC orqali baholanadi [3]. Ushbu yondashuv modelning umumiy diagnostik qobiliyatini to'liq aks ettirmaydi, chunki har bir metrika modelning faqat muayyan jihatini ifodalaydi. Natijada klinik jihatdan muvozanatsiz, ya'ni ayrim muhim holatlarni yetarlicha aniqlay olmaydigan modellar tanlanishi mumkin [4].

Shu sababli ushbu ishda ko'p mezonli baholashga asoslangan integrallashgan yondashuv taklif etiladi. Ushbu yondashuv turli mashinali o'rganish modellarini, jumladan Random Forest, Support Vector Machine va Convolutional Neural Network modellarini yagona mezon asosida kompleks solishtirish imkonini beradi [5–7].

Taklif etilgan metod multi-criteria decision making (MCDM) konsepsiyasiga asoslanib, turli baholash mezonlarini yagona integral ko'rsatkichga birlashtirish orqali optimal modelni tanlashga xizmat qiladi [8]. Shu bilan birga, yondashuv ansambl g'oyalari bilan ham bog'liq bo'lib, bunda bir nechta metrikalarning kombinatsiyasi orqali baholashning barqarorligi oshiriladi [9].

Mazkur yondashuv, shuningdek, turli turdagi ma'lumotlar (tasviriy, klinik va kognitiv) bilan ishlash imkonini beruvchi multi-modal tizimlarda qo'llanilishi

mumkin [10]. Natijada model tanlash jarayoni bitta mezonga bog‘liq bo‘lmay, balki kompleks va klinik jihatdan asoslangan qaror qabul qilishga yo‘naltiriladi.

2. Taklif etilgan metodning matematik ifodasi

Modellar samaradorligini kompleks baholash maqsadida bir nechta metrikalarni yagona integral ko‘rsatkichga birlashtiruvchi funksional yondashuv taklif etiladi:

bu yerda:

- aniqlik (accuracy),
- sezgirlik (sensitivity),
- spetsifiklik (specificity),
- ROC egri chizig‘i ostidagi maydon,
- mos vazn koeffitsientlari.

Vazn koeffitsientlari quyidagi shartni qanoatlantiradi:

Asosiy baholash metrikalari quyidagicha aniqlanadi:

bu yerda α , β , γ , mos ravishda to‘g‘ri va noto‘g‘ri klassifikatsiya natijalarini ifodalaydi.

Taklif etilgan integral ko‘rsatkich model samaradorligini bir nechta mezonlar asosida baholash imkonini berib, eng muvozanatli va klinik jihatdan ishonchli modelni tanlashga xizmat qiladi.

3. Ma‘lumotlar bazasi

Mazkur tadqiqotda Altsgeymer kasalligini erta aniqlash masalasini yechish uchun Alzheimer’s Disease Neuroimaging Initiative (ADNI) ma‘lumotlar bazasidan foydalanildi. Ushbu dataset tibbiy tadqiqotlarda keng qo‘llanilib, multimodal ma‘lumotlarni o‘z ichiga oladi [1].

Tadqiqot doirasida quyidagi ma‘lumot turlaridan foydalanildi:

- MRT tasvirlari (miya tuzilmasini tahlil qilish uchun)
- Klinik va demografik ko‘rsatkichlar
- Kognitiv test natijalari (MMSE va boshqalar)

Obyektlar uchta asosiy sinfga ajratildi: sog‘lom (CN), yengil kognitiv buzilish (MCI) va Altsgeymer kasalligi (AD).

Ma'lumotlar trening va test to'plamlariga mos ravishda 80/20 nisbatda ajratildi hamda model barqarorligini ta'minlash uchun k-fold cross-validation usuli qo'llanildi [2]. Sinflar o'rtasidagi nomutanosiblikni kamaytirish uchun balanslash usullaridan foydalanildi.

Tasvirli ma'lumotlarni tahlil qilishda Convolutional Neural Network asosidagi yondashuv qo'llanildi, jadval ko'rinishidagi ma'lumotlar esa klassik mashinali o'rganish modellariga uzatildi.

4. Ma'lumotlarga ishlov berish (Preprocessing)

Ma'lumotlarga ishlov berish (preprocessing) — bu mashinali o'rganish modeliga uzatilishidan oldin dastlabki ma'lumotlarni tayyorlash bosqichi bo'lib, uning asosiy maqsadi ma'lumotlar sifatini yaxshilash, keraksiz yoki noto'liq qiymatlarni bartaraf etish hamda modelning barqaror va aniq ishlashini ta'minlashdan iborat. Tibbiy diagnostika masalalarida preprocessing ayniqsa muhim, chunki dastlabki ma'lumotlarning sifati model natijalarining ishonchliligiga bevosita ta'sir ko'rsatadi.

Mazkur tadqiqotda jadval ko'rinishidagi ma'lumotlar uchun yetishmayotgan qiymatlar to'ldirildi, xususiyatlar normalizatsiya qilindi va muhim belgilar ajratib olindi. MRT tasvirlari uchun esa o'lchamlarni bir xil formatga keltirish, shovqinni kamaytirish va tasvir sifatini yaxshilash amallari bajarildi. Bundan tashqari, modelning umumlashtirish qobiliyatini oshirish maqsadida data augmentation usullaridan foydalanildi. Ushbu bosqichlar modelni o'qitish jarayonining barqarorligini oshirish va natijalar aniqligini yaxshilashga xizmat qiladi.

5. Natijalar (Results)

Tadqiqot doirasida turli mashinali o'rganish modellarining samaradorligi baholandi. Xususan, Random Forest, Support Vector Machine va Convolutional Neural Network modellarining ishlash natijalari taqqoslandi.

Har bir model uchun asosiy baholash metrikalari — aniqlik (Accuracy), sezgirlik (Sensitivity), spetsifiklik (Specificity) va ROC-AUC hisoblandi. Shuningdek, taklif etilgan integrallashgan baholash ko'rsatkichi qiymatlari aniqlanib, modellarni kompleks baholash amalga oshirildi.

Quyidagi jadvalda modellar samaradorligi keltirilgan:

Model	Accuracy	Sensitivity	Specificity	AUC	Q
Random Forest	0.91	0.88	0.92	0.93	0.91
SVM	0.89	0.85	0.90	0.91	0.89

CNN	0.94	0.92	0.93	0.95	0.94
-----	------	------	------	------	------

Natijalardan ko‘rinib turibdiki, CNN modeli barcha metrikalar bo‘yicha eng yuqori natijalarni ko‘rsatdi. Shu bilan birga, taklif etilgan ko‘rsatkichi modellarni muvozanatli baholash imkonini berdi.

Alohida metrikalar asosida qaralganda, ayrim modellar faqat bitta ko‘rsatkich bo‘yicha ustunlik qilishi kuzatildi. Masalan, Random Forest modeli yuqori spetsifiklikni ta‘minlasa, CNN modeli sezgirlik va AUC bo‘yicha ustunlik ko‘rsatdi.

Taklif etilgan integrallashgan baholash funksiyasi yordamida esa modellar umumiy samaradorlik nuqtai nazaridan baholanib, eng optimal modelni aniqlash imkoniyati yaratildi.

6. Muhokama (Discussion)

Olingan natijalar mashinali o‘rganish algoritmlarining Altsgeymer kasalligini erta aniqlashda yuqori samaradorlikka ega ekanligini tasdiqlaydi. Xususan, Convolutional Neural Network modeli MRT tasvirlaridan xususiyatlarni avtomatik ajratish qobiliyati tufayli boshqa modellar bilan solishtirganda yuqori natijalarni ko‘rsatdi.

Shu bilan birga, Random Forest modeli ham barqaror va ishonchli natijalar berib, ayniqsa klinik ma‘lumotlar bilan ishlashda samarali ekanligi kuzatildi. Support Vector Machine esa kichik o‘lchamli ma‘lumotlarda yaxshi natija ko‘rsatgan bo‘lsa-da, murakkab tasviriy ma‘lumotlar uchun nisbatan pastroq samaradorlikka ega bo‘ldi.

Taklif etilgan integrallashgan baholash yondashuvi alohida metrikalarning cheklovlarini bartaraf etib, modellarni kompleks va ob‘ektiv solishtirish imkonini berdi. Bu yondashuv ayniqsa tibbiy diagnostika masalalarida muhim ahamiyatga ega, chunki noto‘g‘ri qarorlar klinik xatolarga olib kelishi mumkin.

Natijalar shuni ko‘rsatadiki, model tanlash jarayonida faqat bitta metrikaga tayanish yetarli emas, balki bir nechta mezonlarni birgalikda hisobga olish zarur. Shu nuqtai nazardan, taklif etilgan metod klinik qaror qabul qilish tizimlarini takomillashtirishda samarali vosita sifatida qaralishi mumkin.

Xulosa

Mazkur tadqiqotda Altsgeymer kasalligini erta aniqlashda mashinali o‘rganish modellarining samaradorligini baholash masalasi ko‘rib chiqildi. Natijalar shuni ko‘rsatdiki, Convolutional Neural Network modeli tasviriy ma‘lumotlar asosida eng yuqori aniqlikni ta‘minlaydi, Random Forest esa barqaror natijalar ko‘rsatadi.

Taklif etilgan integrallashgan baholash yondashuvi modellarni kompleks va muvozanatli baholash imkonini berib, eng optimal modelni tanlashga yordam

beradi. Ushbu yondashuv tibbiy diagnostika tizimlarida qaror qabul qilish jarayonini takomillashtirishda muhim ahamiyatga ega.

Foydalanilgan adabiyotlar

1. Mueller, S.G., Weiner, M.W., Thal, L.J., et al. (2005). The Alzheimer's Disease Neuroimaging Initiative. *Neuroimaging Clinics of North America*, 15(4), 869–877. <https://doi.org/10.1016/j.nic.2005.09.008>
2. Jack, C.R., Bennett, D.A., Blennow, K., et al. (2018). NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia*, 14(4), 535–562. <https://doi.org/10.1016/j.jalz.2018.02.018>
3. Saito, T., Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE*, 10(3), e0118432. <https://doi.org/10.1371/journal.pone.0118432>
4. Powers, D.M.W. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness and correlation. *Journal of Machine Learning Technologies*, 2(1), 37–63.
5. Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
6. Noble, W.S. (2006). What is a support vector machine? *Nature Biotechnology*, 24, 1565–1567. [<https://doi.org/10.1038/nbt1206-1565>]
7. LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. *Nature*, 521, 436–444. <https://doi.org/10.1038/nature14539>
8. Triantaphyllou, E. (2000). *Multi-Criteria Decision Making Methods: A Comparative Study*. Springer. <https://doi.org/10.1007/978-1-4757-3157-6>
9. Dietterich, T.G. (2000). Ensemble methods in machine learning. *Lecture Notes in Computer Science*, 1857, 1–15. https://doi.org/10.1007/3-540-45014-9_1

10. Baltrušaitis, T., Ahuja, C., Morency, L.P. (2019). Multimodal machine learning: A survey and taxonomy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(2), 423–443. <https://doi.org/10.1109/TPAMI.2018.2798607>
11. Snoek, J., Larochelle, H., Adams, R.P. (2012). Practical Bayesian optimization of machine learning algorithms. *Advances in Neural Information Processing Systems*.