

## YUQORI O'LCHAMLI MA'LUMOTLARDA MODEL MURAKKABLIGI VA BASHORAT ANIQLIGI O'RTASIDAGI MUVOZANAT: SUN'IY INTELLEKT NUQTAYI NAZARIDAN NAZARIY TAHLIL

Xamdamova Fotima G'anijonovna

Toshkent Turizm va Ma'daniy meros texnikumi o'qituvchisi

<https://doi.org/10.5281/zenodo.21945737>

**Annotatsiya:** Mazkur maqolada yuqori o'lchamli ma'lumotlarni tahlil qilishda sun'iy intellekt modellarining murakkabligi va bashorat aniqligi o'rtasidagi nazariy muvozanat masalasi tahlil qilinadi. Tadqiqotda model parametrlarining ko'payishi, ma'lumotlarning o'lchamliligi, ortiqcha moslashuv, umumlashtirish qobiliyati va bashorat xatosi o'rtasidagi o'zaro bog'liqlik tizimlashtiriladi. Klassik bias-variance trade-off konsepsiyasi zamonaviy overparameterized AI modellarining xususiyatlari bilan qiyosiy tahlil qilinib, model murakkabligining ortishi har doim ham umumlashtirish qobiliyatining pasayishiga olib kelmasligi asoslanadi. Shuningdek, regularizatsiya, o'lchamni kamaytirish va model tanlash mexanizmlarining murakkablikni boshqarishdagi roli ochib beriladi. Natijada yuqori o'lchamli muhitda model murakkabligi va bashorat aniqligi o'rtasidagi optimal nisbatni belgilashga qaratilgan nazariy-metodologik yondashuv shakllantiriladi.

**Kalit so'zlar:** yuqori o'lchamli ma'lumotlar; sun'iy intellekt; model murakkabligi; bashorat aniqligi; bias-variance trade-off; overfitting; regularizatsiya; umumlashtirish qobiliyati.

Raqamlashtirish jarayonining jadallashuvi ilmiy tadqiqot, iqtisodiyot, tibbiyot, muhandislik va boshqa ko'plab sohalarda o'zgaruvchilar soni yuqori bo'lgan murakkab ma'lumotlar to'plamlarining shakllanishiga olib kelmoqda. Bunday ma'lumotlarda o'zgaruvchilar sonining kuzatuvlar soniga nisbatan katta bo'lishi tahlil jarayonini an'anaviy statistik muhitdan farqlaydi. Yuqori o'lchamlilik nafaqat hisoblash murakkabligini oshiradi, balki modelning ma'lumotlardagi tasodifiy bog'lanishlarni ham o'zlashtirish ehtimolini kuchaytiradi. Shu sababli yuqori o'lchamli ma'lumotlarni tahlil qilishda modelning qanchalik murakkab bo'lishi kerakligi va ushbu murakkablik bashorat natijasiga qanday ta'sir ko'rsatishi dolzarb metodologik masalaga aylanmoqda.

Sun'iy intellekt texnologiyalarining rivojlanishi mazkur masalaning ahamiyatini yanada oshirdi. Mashinali va chuqur o'rganish modellarining katta hajmdagi parametrlar bilan ishlashi murakkab, nochiziqli va yashirin bog'lanishlarni aniqlash imkonini beradi. Biroq modelning yuqori moslashuvchanligi uning ma'lumotlardagi foydali signal bilan bir qatorda tasodifiy shovqinni ham o'zlashtirishiga olib kelishi mumkin. Natijada training ma'lumotlaridagi yuqori aniqlik yangi ma'lumotlarda xuddi shunday natijaning takrorlanishini kafolatlamaydi. Mazkur holat model murakkabligi va umumlashtirish qobiliyati o'rtasidagi munosabatni nazariy jihatdan qayta ko'rib chiqishni talab qiladi.

An'anaviy statistik o'rganishda ushbu munosabat asosan bias-variance trade-off konsepsiyasi orqali izohlanadi. Modelning haddan tashqari soddaligi ma'lumotlardagi muhim qonuniyatlarni to'liq ifodalay olmaslikka, haddan tashqari murakkabligi esa training ma'lumotlariga ortiqcha moslashishga olib kelishi mumkin. Shuning uchun optimal model soddalik va moslashuvchanlikning muvozanatlashgan holati sifatida qaraladi. Biroq zamonaviy sun'iy intellekt modellarining overparameterized xususiyati ushbu klassik tasavvurni to'liq izohlash imkonini bermaydi. Ayrim hollarda parametrlar sonining keskin ortishiga qaramay,

modelning yangi ma'lumotlarga nisbatan umumlashtirish qobiliyati saqlanib qolishi yoki hatto yaxshilanishi mumkin.

Shu jihatdan model murakkabligini faqat parametrlar soni bilan baholash metodologik jihatdan yetarli emas. Uning samaradorligi ma'lumotlarning o'lchamliligi, signal va shovqin nisbati, modelning arxitekturasi, regularizatsiya darajasi, o'qitish strategiyasi va umumlashtirish qobiliyati bilan o'zaro bog'liq holda namoyon bo'ladi. Ayniqsa, yuqori o'lchamli muhitda keraksiz murakkablikni kamaytirish bilan foydali murakkablikni saqlab qolish o'rtasidagi nisbatni aniqlash muhimdir.

Yuqori o'lchamli ma'lumotlarning asosiy xususiyati o'zgaruvchilar sonining katta bo'lishidir. Bunday sharoitda model ma'lumotlardan ko'proq potensial signal olish imkoniyatiga ega bo'ladi, biroq shu bilan birga ortiqcha va tasodifiy bog'lanishlarni ham o'zlashtirish xavfi kuchayadi. Demak, o'lchamlilikning ortishi modelning imkoniyatlarini kengaytirish bilan birga, uni boshqarish muammosini ham keskinlashtiradi.

Model murakkabligi faqat parametrlar sonini anglatmaydi. U modelning ma'lumotlarni ifodalash qobiliyati, arxitektura chuqurligi, o'zgaruvchilar bilan ishlash usuli, nochiziqli bog'lanishlarni aks ettirish imkoniyati va moslashuvchanligi bilan belgilanadi. Shu sababli bir xil parametrlar soniga ega modellar ham ma'lumotlarni turlicha o'zlashtirishi va turli darajadagi umumlashtirish qobiliyatini namoyon qilishi mumkin.

Yuqori o'lchamli muhitda asosiy metodologik muammo "eng murakkab modelni tanlash" emas, balki **ma'lumotdagi foydali signalni yetarli darajada ifodalay oladigan, biroq tasodifiy shovqinga ortiqcha moslashmaydigan modelni aniqlashdan** iborat. Shu nuqtada model murakkabligi bashorat aniqligidan alohida emas, balki u bilan uzviy bog'liq tushuncha sifatida qaralishi lozim.

Model murakkabligi va bashorat xatosi o'rtasidagi munosabatni tushuntirishda bias-variance trade-off fundamental nazariy tushuncha hisoblanadi. Uning mohiyatiga ko'ra, haddan tashqari sodda model ma'lumotlarning murakkab tuzilmasini yetarli darajada ifodalay olmaydi va yuqori biasga olib keladi. Model murakkablashtirilganda esa uning ma'lumotlarga moslashish qobiliyati ortadi va bias kamayishi mumkin.

Biroq murakkablikning haddan tashqari ortishi modelning muayyan tanlanmaga sezgirligini kuchaytiradi. Bunday model training ma'lumotlaridagi tasodifiy xususiyatlarni ham o'zlashtirib, yangi ma'lumotlarda pastroq natija ko'rsatishi mumkin. Bu holat overfitting bilan izohlanadi. Shunday qilib, klassik yondashuvda bashorat aniqligi modelning soddaligi va murakkabligi o'rtasidagi muvozanatga bog'liq.

Mazkur konsepsiyaning metodologik ahamiyati shundaki, model tanlashda faqat training natijasiga emas, balki uning yangi ma'lumotlarga nisbatan umumlashtirish qobiliyatiga e'tibor qaratish zarurligini asoslaydi. Biroq zamonaviy sun'iy intellekt modellarining katta parametrli tuzilishi klassik bias-variance qarashlarini yangi sharoitda qayta talqin qilishni talab qilmoqda.

Overfitting yuqori o'lchamli ma'lumotlar bilan ishlashda model murakkabligining eng muhim xavflaridan biridir. Model ma'lumotlardagi asosiy signalni emas, balki tanlanmaga xos tasodifiy xususiyatlarni ham o'zlashtirganida uning training natijalari yuqori bo'lishi, ammo yangi ma'lumotlardagi samaradorligi pasayishi mumkin.

Shu bilan birga, zamonaviy chuqur o'rganish modellarida parametrlar soni juda katta bo'lishiga qaramay, yuqori umumlashtirish ko'rsatkichlari kuzatilishi mumkin. Bu holat overparameterization tushunchasini oddiy "ortiqcha murakkablik" sifatida talqin qilishning

yetarli emasligini ko'rsatadi. Modelning katta bo'lishi uning albatta yomon ishlashini anglatmaydi; muhim jihat — ushbu murakkablik qanday boshqarilayotgani va model ma'lumotlarning qaysi xususiyatlarini o'zlashtirayotganidir.

Shu sababli zamonaviy AI nazariyasida model murakkabligi va generalizatsiya o'rtasidagi munosabat bir yo'nalishli emas, balki ko'p omilli jarayon sifatida qaralmoqda. Model arxitekturasini, ma'lumotlar taqsimoti, optimallashtirish va regularizatsiya mexanizmlari birgalikda uning umumlashtirish xususiyatini belgilaydi.

Model murakkabligini boshqarishning asosiy mexanizmlaridan biri regularizatsiyadir. Uning vazifasi modelning ma'lumotlarga moslashuvchanligini nazorat qilish va foydali signal bilan tasodifiy shovqin o'rtasidagi chegarani saqlashdan iborat. Regularizatsiya orqali modelning ortiqcha parametrik moslashuvi cheklanib, umumlashtirish qobiliyatini yaxshilashga sharoit yaratiladi.

Yuqori o'lchamli ma'lumotlarda regularizatsiya ayniqsa muhim, chunki o'zgaruvchilar sonining ko'pligi modelga tasodifiy bog'lanishlarni o'zlashtirish uchun katta imkoniyat beradi. L1 va L2 penalizatsiya, sparsityga asoslangan yondashuvlar, early stopping va boshqa usullar model sig'imini nazorat qilishning turli shakllarini ifodalaydi.

Biroq regularizatsiyani faqat modelni soddalashtirish vositasi sifatida talqin qilish ham to'liq emas. Uning asosiy vazifasi **keraksiz murakkablikni kamaytirish va foydali murakkablikni saqlab qolishdan** iborat. Shu jihatdan regularizatsiya model murakkabligi va bashorat aniqligi o'rtasidagi muvozanatni boshqaruvchi vosita sifatida namoyon bo'ladi.

Yuqori o'lchamli ma'lumotlarda barcha o'zgaruvchilarni modelga bir xil ahamiyatga ega deb kiritish doimo ham samarali emas. Ayrim o'zgaruvchilar takroriy, kuchsiz yoki natija uchun informativ bo'lmagan bo'lishi mumkin. Shu sababli o'lchamni kamaytirish va muhim belgilarni tanlash model murakkabligini nazorat qilishning muhim yo'nalishlaridan hisoblanadi.

O'lchamni kamaytirishning asosiy vazifasi ma'lumot tarkibidagi muhim axborotni saqlagan holda tahlil qilinadigan fazoni ixchamlashtirishdir. Bu nafaqat hisoblash yuklamasini kamaytiradi, balki modelga tasodifiy xususiyatlarga ortiqcha moslashish imkoniyatini ham qisqartirishi mumkin. Biroq o'lchamni kamaytirish jarayonida muhim axborotning yo'qolishi ehtimoli mavjudligi sababli ushbu usul mexanik tarzda emas, ma'lumotlarning strukturaviy xususiyatlari asosida qo'llanishi kerak.

Shundan kelib chiqib, model tanlashda asosiy mezon parametrlar sonining minimal bo'lishi emas, balki **\*\*modelning zarur axborotni yetarli darajada ifodalashi va yangi ma'lumotlarda barqaror natija berishi\*\*** hisoblanadi. Bu yondashuv model soddaligi va bashorat aniqligi o'rtasidagi muvozanatni yanada aniqroq belgilash imkonini beradi.

Sun'iy intellekt modelining haqiqiy samaradorligi uning training ma'lumotlaridagi natijasi bilan emas, balki ilgari ko'rilmagan ma'lumotlarda qanday ishlashi bilan belgilanadi. Shu sababli umumlashtirish qobiliyati model murakkabligini baholashning asosiy mezonlaridan biridir.

Yuqori training aniqligi modelning ma'lumotlarni yaxshi o'rganganini ko'rsatishi mumkin, biroq bu uning yangi ma'lumotlarda ham yuqori natija berishini kafolatlamaydi. Shu sababli modelni baholashda mustaqil test ma'lumotlari, kross-validatsiya va barqarorlik tahlili kabi yondashuvlardan foydalanish muhimdir.

Ayniqsa yuqori o'lchamli ma'lumotlarda modelning umumlashtirish qobiliyati uning parametrlar sonidan ko'ra murakkabroq omillar bilan belgilanadi. Ma'lumotlar sifati, signal va shovqin nisbati, model arxitekturasini, regularizatsiya va o'qitish strategiyasi birgalikda natijaga

taʼsir koʻrsatadi. Shu sababli model murakkabligini bashorat aniqligining yagona determinanti sifatida koʻrish ilmiy jihatdan asoslanmaydi.

Yuqoridagi tahlillar model murakkabligi va bashorat aniqligi oʻrtasidagi munosabatni oddiy “kamroq murakkablik — yaxshiroq model” tamoyili bilan izohlab boʻlmayligini koʻrsatadi. Yuqori oʻlchamli maʼlumotlarda model maʼlum darajadagi murakkablikka ega boʻlishi zarur, chunki haddan tashqari sodda model mavjud signalni toʻliq ifodalay olmaydi. Boshqa tomondan, nazoratsiz murakkablik modelning tasodifiy xususiyatlarga moslashish ehtimolini oshiradi.

Shu sababli optimal muvozanatni **maʼlumotlarning oʻlchamliligi, model sigʻimi, regularizatsiya, umumlashtirish qobiliyati va bashorat barqarorligining** oʻzaro taʼsiri orqali belgilash maqsadga muvofiq. Bunda modelning maksimal murakkabligi emas, balki maʼlumotdan foydali signalni samarali ajratib olish imkonini beradigan murakkablik darajasi ustuvor hisoblanadi.

Zamonaviy AI modellarida kuzatilayotgan overparameterization va double descent hodisalari ham ushbu muvozanatni statik emas, balki dinamik tushuncha sifatida qarash zarurligini koʻrsatadi. Demak, model murakkabligi va bashorat aniqligi oʻrtasidagi optimal nisbat har bir maʼlumotlar muhiti uchun alohida shakllanadi va uni faqat parametrlar soniga asoslangan yagona mezon bilan belgilab boʻlmaydi.

Nazariy tahlil natijasida yuqori oʻlchamli maʼlumotlarda model murakkabligi va bashorat aniqligi oʻrtasidagi munosabat klassik bias–variance trade-off doirasida boshlangʻich izohga ega boʻlsa-da, zamonaviy sunʼiy intellekt modellarining xususiyatlari sababli yanada murakkab tus olishi aniqlandi. Modelning soddaligi yetarli signalni ifodalash imkonini cheklashi, ortiqcha murakkablik esa maʼlumotlardagi tasodifiy xususiyatlarga moslashish xavfini kuchaytirishi mumkin. Shu bilan birga, overparameterized modellar tajribasi murakkablikning oʻzi generalizatsiyaning pasayishini anglatmasligini koʻrsatadi.

Shu asosda birinchi tavsiya — yuqori oʻlchamli maʼlumotlar uchun model tanlashda parametrlar sonini yagona mezon sifatida qabul qilmaslik. Model murakkabligi maʼlumotlarning strukturasi, signal va shovqin nisbati hamda kutilayotgan vazifa bilan bogʻliq holda baholanishi lozim.

Ikkinchi tavsiya — model murakkabligini boshqarishda regularizatsiya, oʻlchamni kamaytirish va muhim belgilarni tanlash mexanizmlaridan maqsadli foydalanish. Bunda maqsad modelni maksimal darajada soddalashtirish emas, balki keraksiz murakkablikni kamaytirishdir.

Uchinchi tavsiya — AI modellarining samaradorligini faqat training aniqligi asosida baholamaslik. Umumlashtirish qobiliyati, mustaqil test natijalari, kross-validatsiya va model barqarorligi birgalikda hisobga olinishi kerak.

Toʻrtinchi tavsiya — overparameterization sharoitida klassik bias–variance trade-offni mutlaq qoida sifatida emas, model murakkabligi va generalizatsiya munosabatini tushuntiruvchi boshlangʻich nazariy paradigma sifatida qoʻllash. Ayniqsa double descent kabi hodisalar mavjud boʻlgan sharoitlarda modelning real xatti-harakati empirik va nazariy jihatdan kompleks baholanishi zarur.

Beshinchi tavsiya — yuqori oʻlchamli maʼlumotlar bilan ishlashda “eng murakkab model” yoki “eng sodda model” tamoyillaridan voz kechib, **maʼlumotga mos murakkablik** tamoyilini asosiy metodologik mezon sifatida qabul qilish. Bunday yondashuv bashorat aniqligi, umumlashtirish va model barqarorligi oʻrtasidagi muvozanatni taʼminlashga xizmat qiladi.

Umuman, yuqori o'lchamli ma'lumotlarda optimal sun'iy intellekt modeli parametrlar soni eng kam yoki eng ko'p bo'lgan model emas, balki **ma'lumotdagi foydali signalni yetarli darajada ifodalash, ortiqcha moslashuvni nazorat qilish va yangi ma'lumotlarda barqaror natija berish imkoniyatiga ega bo'lgan modeldir**. Shu ma'noda model murakkabligini boshqarish sun'iy intellekt tizimlarining faqat texnik emas, balki muhim nazariy-metodologik masalasi sifatida qaralishi lozim.

### **Adabiyotlar, References, Литературы:**

1. Belkin M., Hsu D., Ma S., Mandal S.\*\* Reconciling Modern Machine-Learning Practice and the Classical Bias-Variance Trade-Off // Proceedings of the National Academy of Sciences. — 2019. — Vol. 116, No. 32. — P. 15849–15854.
2. Nakkiran P., Kaplun G., Bansal Y., Yang T., Barak B., Sutskever I.\*\* Deep Double Descent: Where Bigger Models and More Data Hurt // Journal of Statistical Mechanics: Theory and Experiment. — 2021. — 124005.
3. D'Ascoli S., Refinetti M., Biroli G., Krzakala F.\*\* Double Trouble in Double Descent: Bias and Variance(s) in the Lazy Regime // Proceedings of the 37th International Conference on Machine Learning. — 2020. — Vol. 119. — P. 2280–2290.
4. Fan J., Li R. Variable Selection via Nonconcave Penalized Likelihood and its Oracle Properties // Journal of the American Statistical Association. — 2001. — Vol. 96, No. 456. — P. 1348–1360.
5. Bühlmann P. Statistical Significance in High-Dimensional Linear Models // Bernoulli. — 2013. — Vol. 19, No. 4. — P. 1212–1242.
6. Nalsnick E., Smyth P., Tran D. A Brief Tour of Deep Learning from a Statistical Perspective // Annual Review of Statistics and Its Application. — 2023. — Vol. 10. — P. 219–246. — DOI: 10.1146/annurev-statistics-032921-013738.
7. Qizi, Jalilova Dilshoda Ural, Jo'raboyeva Malika, Jumayeva Muhlis, and Jamolova Sevinch. "BOSHLANG 'ICH SINFLAR ELEKTRON TA'LIMDA MULTIMEDIANING O'RNINI." *Science and innovation* 3, no. Special Issue 22 (2024): 206-207.
8. Qizi, Jalilova Dilshoda Ural, Poyonova Qandiya, Abdimurodova Maftuna, and Nahalboyeva Iroda. "UCHINCHI RENESSANS SHAROITIDA PEDAGOGIK TA'LIM SOHASIDA RAQAMLI TEXNOLOGIYALAR." *Science and innovation* 3, no. Special Issue 18 (2024): 637-640.
9. Winkle C.K., Zammit-Mangion A. Statistical Deep Learning for Spatial and Spatiotemporal Data // Annual Review of Statistics and Its Application. — 2023. — Vol. 10. — P. 247–270. — DOI: 10.1146/annurev-statistics-033021-112628.
10. Statistical Inference: Theoretical Development to Data Analytics // Handbook of Statistics. — 2020. — Vol. 43. — P. 289–335. — DOI: 10.1016/bs.host.2020.02.003.