

**SUN'IY INTELLEKT ASOSIDA HUJJAT TAHLIL QILISH:
ZAMONAVIY YONDASHUVLAR**

Abduraxmonova Gulimoh Abdullayevna

Toshkent davlat yuridik universiteti
Xususiy huquq fakulteti 1-kurs talabasi
abduraxmonovagulimoh@icloud.com

Annotatsiya. *Sun'iy neyron tarmoqlar hujjatlarni tahlil qilish va tanib olishda keng qo'llanilgan. Aksariyat sa'y-harakatlar alohida yozilgan va bosma belgilarni tanib olishga qaratilgan bo'lib, bunda sezilarli muvaffaqiyatlarga erishilgan. Biroq, hujjatlarni qayta ishlashning boshqa ko'plab vazifalari, masalan, oldindan qayta ishlash (preprocessing), sahifa tuzilishini tahlil qilish (layout analysis), belgilarni ajratish (character segmentation), so'zlarni tanib olish (word recognition) va imzoning haqiqiylikini tekshirish (signature verification) kabi masalalar ham istiqbolli natijalar bilan samarali hal etilgan. Maqolada sun'iy intellekt asosida hujjatlarni avtomatlashtirilgan tarzda tahlil qilish, strukturalangan va strukturasiz matnli hujjatlar — ustida SI yondashuvlari qo'llanilishi, metodologik jihatlari, zamonaviy platformalar va amaliy natijalari tahlil qilinadi.*

Kalit so'zlar: *sun'iy intellekt (SI), hujjat tahlili, hujjatni tasniflash, NLP, hujjatni ajratib olish, avtomatlashtirish, document AI, tilni qayta ishlash.*

Kirish

Bugungi kunda axborot texnologiyalari jamiyatning ajralmas bir qismiga aylanib borayotgan bir vaziyatda, sun'iy intellekt kabi texnologiyalardagi yutuqlar kundalik hayotimizda shaxsiy hamda ish faoliyatimizni amalga oshirish chog'ida bir qancha o'zgarishlarga olib kelmoqda. XXI asr "raqamli asr" deb atalgan davrda sodir bo'layotgan o'zgarishlar o'z ta'sirini bevosita huquqiy sohaga ham o'tkazmoqda. Raqamli hamda axborot texnologiyalarining jadal rivojlanishi

huquqiy texnologiyalarni ham o'zgartirishni talab qiladi. Huquqiy texnologiyalarni tushunishda yangi yondashuvlar paydo bo'lmoqda, raqamlashtirish jarayonining huquqiy voqelikka ta'sirini hisobga oladigan yangi kategoriyalar ishlab chiqilmoqda.

“Men qog'oz hujjat yoki PDF dagi ma'lumotlarni o'qishga va bu ma'lumotlarni buxgalteriya hisobi dasturiga kiritishga vaqt sarflashni istayman”, deb hech kim aytmaydi. Aynan shuning uchun ham AI asosidagi hujjatlarni tushunish texnologiyasi shunchalik foydali — u odamlar xursandchilik bilan topshiradigan ishlarni o'z zimmasiga oladi. Bu orqali u hujjatlarni qayta ishlashni tezlashtirishi, xarajatlarni kamaytirishga yordam berishi va aniqlikni oshirishi mumkin. Dokumentlarni tushunish texnologiyasi ko'pincha sizning SaaS (xizmat sifatida dasturiy ta'minot) ilovangiz ish oqimlariga kiritilgan bo'ladi, va so'nggi paytlarda u AI o'quv ma'lumotlari uchun matnlarni belgilash va AI agentlari o'zlarining ko'p bosqichli vazifalarini bajarishlari uchun zarur bo'lgan ma'lumotlarni taqdim etish uchun ham ishlatilmoqda. Shu tariqa, odamlar o'zlarining fikrlash, qurish va muloqot qilish kabi yoqtirgan ishlariga qaytishlari mumkin.¹

Hozirgi davrda tashkilotlar va ilmiy muassasalar katta hajmdagi hujjatlar bilan ishlamoqda — kontraktlar, texnik hisobotlar, qonun hujjatlari, arxiv materiallari va h.k. An'anaviy qo'l bilan hujjat tahlili jarayoni ko'p vaqt oladi, xatolik ehtimoli yuqori bo'ladi va samaradorlik cheklangan. Shu nuqtai nazardan, sun'iy intellekt (SI) texnologiyalari — jumladan, tabiiy tilni qayta ishlash (NLP), optik belgilarni tanish (OCR), mashinasozlik tahlili — hujjatlarni avtomatlashtirilgan tarzda tahlil qilish imkoniyatini yaratmoqda. Masalan,

¹ Erickson, J. (2025, July 31). Artificial Intelligence-Based Document Analysis: Modern Approaches. Retrieved from <https://www.addepto.com/blog/how-ai-is-revolutionizing-document-analysis-a-comprehensive-guide>.
SJIF: 5.051

LayoutLMv3 modeli keltirilgan hujjat tasvirlari va matnlarni birgalikda qayta ishlash orqali yaxshi natijalarga erishdi.

So‘nggi yillarda SI va hujjat tahlili sohasida bir qator statistik ko‘rsatkichlar dolzarbligini bir-an oshirib bormoqda. Masalan, global SI bozorining hajmi 2024-yilda taxminan 638 milliard dollarni tashkil etgan va 2034-yilgacha taxminan 3,68 trillion dollargacha o‘sishi kutilmoqda.²

Bundan tashqari, hujjatlarni avtomatlashtirilgan tahlil qilish bo‘yicha — ya’ni “intelligent document processing” (IDP) texnologiyalarini qo‘llash ham sezilarli darajada rivojlanmoqda: global IDP bozorining hajmi 2024-yilda taxminan 2,3 milliard dollar bo‘lgan va 2025–2034 yillarda yiliga $\approx 24,7\%$ CAGR bilan o‘sishi kutilmoqda.³

Ushbu maqolaning maqsadi — SI asosida hujjat tahlili bo‘yicha zamonaviy yondashuvlarni tahlil qilish, metodologiyalarni sharhlash, real dunyo tadqiqotlari va natijalarini ko‘rib chiqish, hamda kelgusida qo‘llash bo‘yicha tavsiyalar berishdir.

Material va metodlar

Mazkur ilmiy maqolani tayyorlashda tarixiy, tizimli-tuzilmaviy, mantiqiy, induksiya, deduksiya, qiyosiy-huquqiy, ilmiy manbalarni kompleks tadqiq etish va statistik ma’lumotlar tahlil qilish kabi usullardan foydalanilgan.

Tadqiqotda adabiyotlar tahlili asos sifatida foydalanildi. Xususan: SI hujjat tahlili bo‘yicha so‘nggi xalqaro maqolalar ko‘rib chiqildi va NLP (Natural Language Processing) metodlari, RAG (Retrieval-Augmented Generation)

² Global Market Insights (GMI). (2024). Intelligent Document Processing (IDP) Market Size Report, 2024–2034. Havola: <https://www.gminsights.com/industry-analysis/intelligent-document-processing-market>.

³ Precedence Research. (2024). Artificial Intelligence (AI) Market Size, Growth and Forecast 2024–2034. Havola: <https://www.precedenceresearch.com/artificial-intelligence-market>.

yondashuvi, Information Extraction (IE), OCR (Optical Character Recognition), Machine Learning va Deep Learning modellaridan foydalanildi.

Tadqiqot natijalari

Dokumentlarni tushunish – bu PDF yoki qog‘oz hujjat skaneri kabi matn faylidan ma’lumotlarni ajratib oladigan va ularni biznes dasturiga uzatadigan avtomatlashtirilgan jarayon. Bu jarayon qo‘lda ma’lumot kiritishni minimallashtirishga va hatto butunlay yo‘q qilishga yordam beradi, shu bilan birga aniqlikni oshiradi. Dokumentlarni tushunish murakkab mashina o‘rganish (MO‘) algoritmlari tufayli amalga oshiriladi. MO‘ bir necha bosqichda, jumladan, tegishli ma’lumotlar, masalan, narx, ism va hisob-faktura yoki xarid buyurtmasi raqami aniqlanadigan, ajratib olinadigan va tegishli biznes tizimlariga integratsiya qilish uchun ma’lumotlar bazasida saqlanadigan tasvirni qayta ishlashdan boshlab, muhim rol o‘ynaydi. Masalan, bir ishlab chiqaruvchi xarid buyurtmalaridan ma’lumotlarni ajratib olish va ularni avtomatik tarzda buxgalteriya kitobiga va inventar nazorati tizimiga kiritish uchun hujjatlarni tushunishdan foydalanishi mumkin, bu esa savdo jarayonining tezligi va aniqligini sezilarli darajada oshiradi. Shuningdek, bir kompaniya kvitansiya rasmlaridan tegishli ma’lumotlarni olib, xodim uchun avtomatik ravishda xarajat hisobotini yaratadigan xarajatlarni qoplash tizimini o‘rnatishi mumkin.

Dokumentlarni tushunish AI va MO‘ ning erta va muvaffaqiyatli qo‘llanilishi bo‘lgan. U biznes ilovalariga integratsiya qilinadi va ish oqimlarini avtomatlashtirishga yordam beradi, bu esa qo‘lda ma’lumot kiritish va hujjatlarni qayta ishlash bilan shug‘ullanadigan mutaxassislarning vaqtini tejaydi. Tabiiy tilni qayta ishlash (NLP) va retrieval-augmented generation (RAG) bilan birgalikda, dokumentlarni tushunish hujjatlarning semantik ma’nosini tushunishga yordam

beradigan tizimning ajralmas qismiga aylanadi va hujjatlarni tasniflash va ma'lumotlarni aniqlashda yordam beradi.⁴

Xuddi inson miyasi ma'lumotlarni qayta ishlaganidek, Document AI dagi neyron tarmoqlar hujjatlarning tuzilishi, shriftlari va tillarini tahlil qilib, turli formatlarga uzluksiz moslashadi. Bu moslashuvchanlik Document AI ga oddiy hisob-fakturalardan tortib, murakkab yuridik shartnomalargacha bo'lgan ko'plab real hayotiy ssenariylarni tahlil qilish va doimiy o'rganish orqali o'z imkoniyatlarini yaxshilash imkonini beradi.

Shuningdek, metadata ham hujjat haqida qo'shimcha, ko'pincha yashirin ma'lumotlarni taqdim etish orqali muhim rol o'ynaydi. Metadata tarkibiga hujjatning yaratilgan sanasi, muallifi, fayl formati va uning mazmunini yanada aniqroq tavsiflovchi kalit so'zlar kabi ma'lumotlar kiradi. Metadatadan foydalanish orqali Document AI hujjatlarni yaxshiroq tashkil qilish, boshqarish va qidirish uchun harakat qiladi, shu bilan ish oqimining samaradorligini oshiradi.

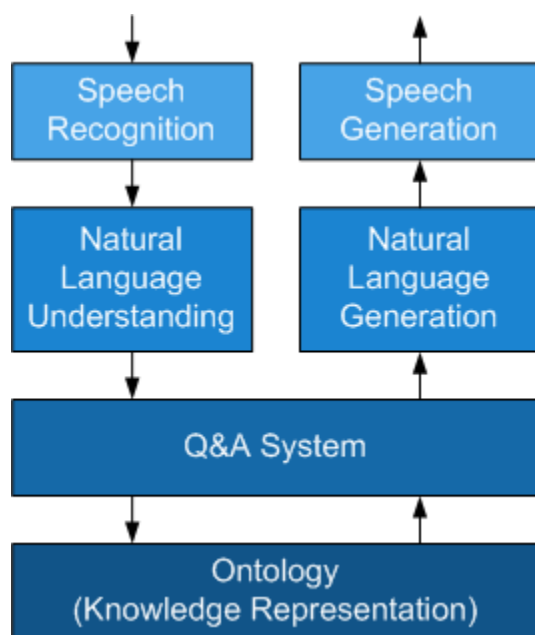
1954-yilda IBM kompaniyasi IBM 701 meynfreymida mashina tarjimasidan foydalangan holda ruscha gaplarni ingliz tiliga tarjima qilish imkoniyatini namoyish etdi. Bugungi standartlarga ko'ra oddiy bo'lishiga qaramay, bu namoyish tilni tarjima qilishning ulkan afzalliklarini ko'rsatib berdi. Ushbu maqolada biz tabiiy tilni qayta ishlash (NLP) va uning kompyuterlar bilan yanada tabiiyroq muloqot qilishimizga qanday yordam berishini ko'rib chiqamiz.

NLP bir qancha sabablarga ko'ra mashina o'rganishning eng muhim kichik yo'nalishlaridan biridir. Tabiiy til foydalanuvchi va mashina o'rtasidagi eng tabiiy interfeys hisoblanadi. Ideal holatda, bu nutqni aniqlash va ovoz yaratishni o'z ichiga oladi. Hatto Alan Tyuring ham buni o'zining "intellekt" haqidagi

⁴ Erickson, J. (2025, July 31). Artificial Intelligence-Based Document Analysis: Modern Approaches. Retrieved from <https://www.addepto.com/blog/how-ai-is-revolutionizing-document-analysis-a-comprehensive-guide>.
SJIF: 5.051

maqolasida tan olgan, unda u "Tyuring testi"ni tabiiy tildagi suhbat orqali mashinaning aqli xatti-harakatlarni namoyish etish qobiliyatini tekshirish usuli sifatida belgilagan.

NLP bitta alohida mavjudot emas, balki tadqiqot yoʻnalishlari spektridir. 1-rasmda bugungi kunda NLPning keng tarqalgan mahsuloti boʻlgan ovozli yordamchi koʻrsatilgan. NLP tadqiqot yoʻnalishlari ovozli yordamchi dasturining asosiy bloklari kontekstida koʻrsatilgan.⁵



Ovozli yordamchi texnologiyasidan tashqari, NLPning yana bir asosiy afzalligi shundaki, dunyoda juda koʻp miqdordagi tartibsiz matnli maʼlumotlar mavjud boʻlib, bu maʼlumotlar tabiiy tilni qayta ishlash va tushunish uchun asosiy harakatlantiruvchi kuch boʻlib xizmat qiladi. Mashina ushbu matnni (asosan, inson isteʼmoli uchun yaratilgan) qayta ishlashi, tartibga solishi va tushunishi orqali biz kelajakdagi mashina oʻrganish ilovalari uchun koʻplab foydali dasturlarni va ishga

⁵ Jones, M. T. (2025). A beginner's guide to natural language processing: Discover how natural language processing can help you to converse more naturally with computers. IBM Developer. Retrieved from <https://developer.ibm.com/articles/a-beginners-guide-to-natural-language-processing/>.

solinadigan ulkan miqdordagi bilimlarni ochishimiz mumkin. Vikipediya, masalan, mavzular o'rtasidagi munosabatlarni ko'rsatish uchun ko'p jihatdan bog'langan katta hajmdagi bilimlarni o'z ichiga oladi. Vikipediya o'z-o'zidan NLPni qo'llash mumkin bo'lgan bebaho tartibsiz ma'lumot manbai bo'ladi.

NLPga zamonaviy yondashuvlar birinchi navbatda neyron tarmoq arxitekturalariga e'tibor qaratadi. Neyron tarmoq arxitekturalari raqamli ishlov berishga tayanganligi sababli, so'zlarni qayta ishlash uchun kodlash talab qilinadi. Ikkita keng tarqalgan usul - bir issiq kodlash va so'z vektorlari.

Bir issiq kodlash (one-hot encoding) — bu so'zlarni raqamli shaklga o'tkazish usuli bo'lib, har bir so'z noyob vektor bilan ifodalanadi. Ya'ni, lug'atdagi har bir so'z uchun uzunligi so'zlar soniga teng bo'lgan vektor yaratiladi va unda faqat bitta element "1", qolganlari esa "0" bo'ladi. Bu usul neyron tarmoqlarga so'zlarni raqam sifatida kiritish imkonini beradi. Natijada, tarmoq har bir so'zni osonroq ajratib, samaraliroq o'qitiladi.⁶

the	1	0	0	0	0
man	0	1	0	0	0
we	0	0	1	0	0
saw	0	0	0	1	0
a	0	0	0	0	1

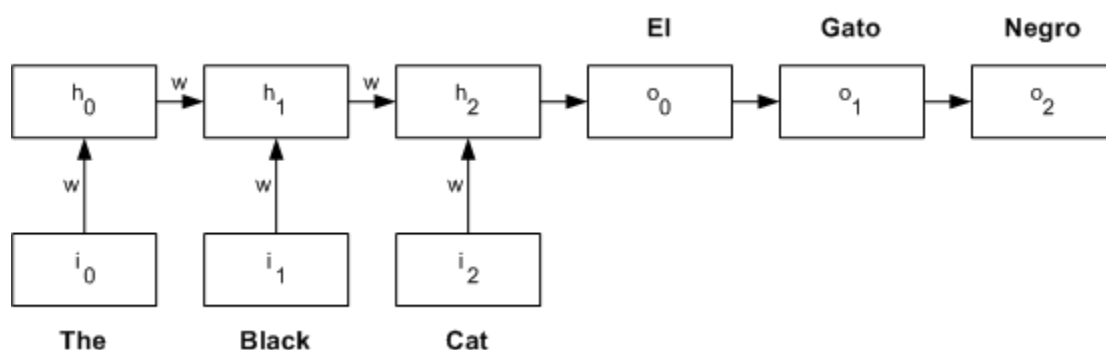
Yana bir kodlash - bu so'z vektorlari, ular so'zlarni yuqori o'lchamli vektorlar sifatida ifodalaydi, bu yerda vektorning birliklari haqiqiy qiymatlardir. Ammo har bir birlikga so'zni tayinlash o'rniga (masalan, bir issiq kabi), har bir birlik so'z

⁶ Jones, M. T. (2025). A beginner's guide to natural language processing: Discover how natural language processing can help you to converse more naturally with computers. IBM Developer. Retrieved from <https://developer.ibm.com/articles/a-beginners-guide-to-natural-language-processing/>.

uchun toifalarni ifodalaydi (masalan, yakka va ko'plik yoki ot va fe'l) va 100-1,000 birlik kengligida bo'lishi mumkin (o'lchamlilik). Ushbu kodlashni qiziqarli qiladigan narsa shundaki, so'zlar endi raqamli bog'liqdir va kodlash matematik operatsiyalarni so'z vektorlariga (masalan, qo'shish, olib tashlash yoki rad etish kabi) qo'llashni ochadi.

So'z vektorlari kodlash — bu so'zlarni haqiqiy sonlardan tashkil topgan uzun vektorlar yordamida ifodalash usulidir. Bunda har bir son (yoki birlik) so'zning ma'nosi yoki grammatik xususiyatini bildiradi. Masalan, so'zning yakka yoki ko'plik shakli, ot yoki fe'l bo'lishi kabi. Bu vektorlar odatda 100 dan 1000 gacha o'lchamga ega bo'ladi. Eng muhim tomoni — bu usulda so'zlar o'zaro ma'noga yaqin bo'lsa, ularning vektorlari ham bir-biriga o'xshash bo'ladi. Shu tufayli ularni matematik tarzda solishtirish yoki tahlil qilish mumkin.

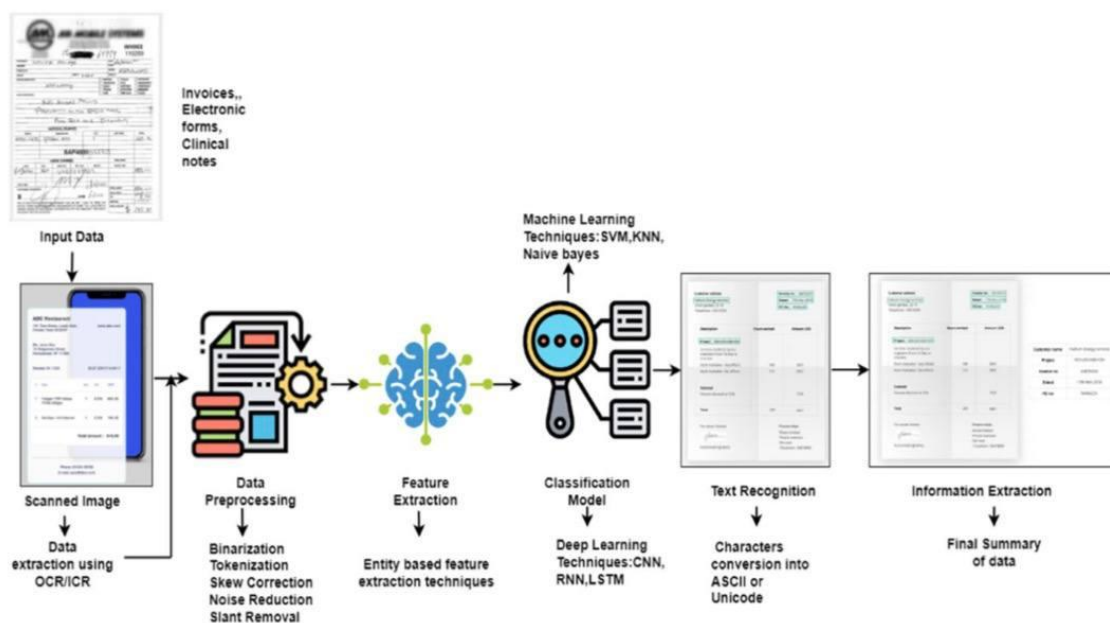
1980-yillarda yaratilgan takroriy neyron tarmoqlar (RNN) tabiiy tilni qayta ishlashda (NLP) muhim rol o'ynaydi. Oddiy neyron tarmoqlardan farqli ravishda, RNNlar ma'lumotlarni vaqt ketma-ketligi bo'yicha tahlil qiladi. Ya'ni, ular har bir bosqichda oldingi natijani eslab qoladi va keyingi bosqichda undan foydalanadi. Shu sababli RNN matn yoki nutq kabi ketma-ket ma'lumotlar bilan ishlashga juda qulay. Masalan, RNN so'zlarni nafaqat alohida holda, balki ular joylashgan kontekst bilan birgalikda o'rganadi — ya'ni, oldin va keyin kelgan so'zlarni ham hisobga oladi. Amalda esa, tarmoq har bir so'zning kodini ketma-ket qabul qiladi va ularni bosqichma-bosqich qayta ishlaydi. Oxirida esa, barcha so'zlar kiritilgandan keyin yakuniy natija (masalan, tarjima qilingan so'zlar) hosil bo'ladi.



Tadqiqot natijalar tahlili

Raqamli matnni tanish texnologiyasi ilmiy tadqiqotlar sohasida tez sur'atlar bilan rivojlanmoqda. Matnni tanish tizimi inson va kompyuter o'rtasida aloqa o'rnatishga yordam beradigan interfeys yaratishda qo'llaniladi. Chop etilgan yoki qo'lda yozilgan matnlar raqamli shaklda kodlanadi va matnni tanish jarayonida ishlatiladi. Hozirgi tadqiqotlarni yaxshiroq tushunish uchun ushbu bo'limda tegishli ilmiy adabiyotlarning batafsil sifat tahlili keltirilgan. Tahlil davomida "(unstructured text-based AND data analysis AND data recognition or e-governance or clinical notes)" kalit so'zlari orqali tuzilmasiz hujjatlarda chop etilgan yoki qo'lda yozilgan matnni tanish bo'yicha tadqiqotlarning ta'siri o'rganildi. Tanlangan maqolalar ko'rib chiqilgach, eng moslari chuqurroq tahlil uchun ajratib olindi. Shuningdek, ushbu tadqiqotda tuzilmasiz hujjatlardagi ma'lumotlarni aniqlash va xulosalash jarayoni sog'liqni saqlash, sug'urta, elektron boshqaruv (e-gov) kabi turli sohalarda qaror qabul qilish jarayoniga qanday yordam berishi mumkinligi o'rganildi. Qisqacha ma'lumotlarni ajratib olish muammosini hal etishga bo'lgan qiziqish mashinaviy o'rganish qo'llaniladigan sohalarda — masalan, Kompyuter ko'rish (Computer Vision, CV) va Tabiiy tilni qayta ishlash (Natural Language Processing, NLP) yo'nalishlarida ortib bormoqda. Tuzilmasiz hujjatlarning shakliga qarab, eng samarali natijaga erishish uchun turli

sun'iy intellekt (AI) asosidagi yondashuvlardan foydalaniladi. Bugungi kunda internet va ijtimoiy tarmoqlarda ulkan hajmdagi tuzilmasiz ma'lumotlar oqimi kuzatilmoqda. Shu sababli foydalanuvchilarga bu katta hajmdagi tartibsiz ma'lumotlar orasidan foydali ma'lumotni tezda topish qiyin bo'lmoqda. Taklif etilgan tuzilmasiz hujjatlarni qayta ishlash tizimi asosan to'rt qismga bo'lingan bo'lib, bu jarayon 3-rasmda (yoki keyingi rasmda) tasvirlangan.⁷



Tuzilmagan (ya'ni tartibga solinmagan) hujjatlardan, masalan, skaner qilingan hujjatlardan matnni qo'lda ajratib olish mehnat talab qiladi va inson omili tufayli xatolarga moyil bo'ladi. Shu sababli tashkilotlar hujjatlarni qayta ishlash jarayonini avtomatlashtirish uchun shablon asosidagi usullar, masalan, OCR (Optical Character Recognition — optik belgilarni tanish) texnologiyasidan foydalanishga harakat qilgan. OCR texnologiyasi tasvir ichida joylashgan matnni

⁷ Mahadevkar, S. V., Patil, S., Kotecha, K., Lim, W. S., & Choudhury, T. (2024). Exploring AI-driven approaches for unstructured document analysis and future horizons. *Journal of Big Data*, 11(92). <https://doi.org/10.1186/s40537-024-00948-z>
SJIF: 5.051

aniqlash uchun qo'llaniladi — bu ko'pincha bosma yoki qo'lda yozilgan hujjatning skaner qilingan nusxasi bo'ladi. OCR yordamida turli xil hujjat turlarini avtomatik ravishda toifalarga ajratish va belgilangan qoidalarga muvofiq tartiblash mumkin. Masalan, shartnomalarni sotuvchi yoki mahsulot turi asosida saqlab, yangilab turish uchun foydalaniladi.

Tahlil davomida quyidagi umumiy kuzatuvlar aniqlandi:

- AI tizimlari bosma matn bilan ishlaganda juda aniq natija beradi, ammo qo'lyozma matnlarda hali ham xatoliklar mavjud.
- Ko'p hollarda ma'lumotlarning sifatli tayyorlanmaganligi (masalan, shovqinli yoki notekis skaner) model aniqligini pasaytirgan.
- Ko'p modal yondashuvlar (matn + tasvir) yagona modal yondashuvlarga qaraganda sezilarli darajada aniqroq natija bergan.
- Transformer oilasiga mansub modellar (BERT, GPT, LayoutLM) so'nggi yillarda eng samarali deb topilgan.

Asosiy xulosalar quyidagicha:

- Deep Learning modellar (xususan Transformer va CNN) eng yuqori aniqlikka ega.
- Tibbiyot va moliya sohalari AI asosidagi hujjat tahlilini eng faol joriy qilgan yo'nalishlar.
- Mavjud datasetlar cheklangan, shu bois yangi, murakkab va real hujjatlarni o'z ichiga olgan ma'lumot to'plamlari kerak.
- AI tizimlari an'anaviy OCR usullaridan ancha ustun — xatolikni kamaytiradi va tahlil tezligini oshiradi.

Xulosa

Bugungi kunda sun'iy intellekt (AI) texnologiyalarining jadal rivojlanishi hujjatlarni avtomatik tahlil qilish jarayonida inqilobiy o'zgarishlarni yuzaga keltirmoqda. An'anaviy qo'lda bajariladigan yoki shablon asosidagi tizimlardan farqli ravishda, AI texnologiyalari — xususan Deep Learning, Natural Language Processing (NLP) va Optical Character Recognition (OCR) usullari — tuzilmagan hujjatlardan ma'lumotlarni aniqlik bilan ajratib olish, tasniflash va raqamli shaklga keltirish imkonini bermoqda. So'nggi yillarda Transformer va konvolyutsion neyron tarmoqlar (CNN) asosidagi modellar yuqori aniqlik va moslashuvchanlikni ta'minlab, turli sohalarda — tibbiyot, moliya, huquq, ta'lim va davlat boshqaruvida — muvaffaqiyatli qo'llanilmoqda. Shu bilan birga, hali ham sifatli ma'lumotlar bazalarining yetishmasligi, ma'lumotlarning maxfiyligi hamda modelning umumlashtirish qobiliyatidagi cheklovlar ushbu yo'nalishda dolzarb muammo bo'lib qolmoqda.

Kelajakda AI tizimlarini yanada takomillashtirish uchun quyidagi yo'nalishlarga e'tibor qaratish zarur:

- Ko'p tilli va real hayotiy ma'lumotlarga asoslangan datasetlarni yaratish,
- Izohlanadigan (Explainable) va yengil AI modellarini ishlab chiqish,
- AI, RPA (Robotic Process Automation) va NLP texnologiyalarini birlashtirgan gibrid yondashuvlarni keng joriy etish.

Shunday qilib, sun'iy intellekt asosida hujjatlarni tahlil qilish — nafaqat ma'lumotlarni avtomatlashtirish vositasi, balki zamonaviy boshqaruv tizimlari, raqamli transformatsiya va elektron hukumat infratuzilmasining ajralmas qismi sifatida shakllanmoqda. Bu yo'nalishning izchil rivojlanishi inson mehnatini yengillashtirish, xatoliklarni kamaytirish va samaradorlikni sezilarli darajada oshirishga xizmat qiladi.

REFERENCES

1. Chakrabarti, D., Patodia, N., Bhattacharya, U., Mitra, I., Roy, S., Mandi, J., Roy, N., & Nandy, P. (2018, October). Use of artificial intelligence to analyse risk in legal documents for a better decision support. In Proceedings of the TENCON—IEEE Region 10 Conference (pp. 683–688). IEEE. <https://doi.org/10.1109/TENCON.2018.8650382>.
2. Clausner, C., Antonacopoulos, A., & Pletschacher, S. (2020, March). Efficient and effective OCR engine training. International Journal on Document Analysis and Recognition, 23(1), 73–88. <https://doi.org/10.1007/s10032-019-00347-8>.
3. Erickson, J. (2025, July 31). Artificial intelligence-based document analysis: Modern approaches. Addepto. <https://www.addepto.com/blog/how-ai-is-revolutionizing-document-analysis-a-comprehensive-guide>.
4. Global Market Insights. (2024). Intelligent document processing (IDP) market size report, 2024–2034. <https://www.gminsights.com/industry-analysis/intelligent-document-processing-market>.
5. Jones, M. T. (2025). A beginner's guide to natural language processing: Discover how natural language processing can help you to converse more naturally with computers. IBM Developer. <https://developer.ibm.com/articles/a-beginners-guide-to-natural-language-processing/>.
6. Mahadevkar, S. V., Patil, S., Kotecha, K., Lim, W. S., & Choudhury, T. (2024). Exploring AI-driven approaches for unstructured document analysis and

future horizons. Journal of Big Data, 11(92). <https://doi.org/10.1186/s40537-024-00948-z>.

7. Precedence Research. (2024). Artificial intelligence (AI) market size, growth and forecast 2024–2034. <https://www.precedenceresearch.com/artificial-intelligence-market>.

8. Singh, P., Varadarajan, S., Singh, A. N., & Srivastava, M. M. (2020). Multi-domain document layout understanding using few-shot object detection. In Image Analysis and Recognition (Lecture Notes in Computer Science, Lecture Notes in Artificial Intelligence, and Lecture Notes in Bioinformatics, Vol. 12132, pp. 89–99). Springer.

9. Zaman, G., Mahdin, H., & Hussain, K. (2020, June). Information extraction from semi and unstructured data sources: A systematic literature review. ICIC Express Letters, 14(6), 593–603. <https://doi.org/10.24507/icicel.14.06.593>.