MA'LUMOTLAR BILAN ISHLASHDA SUN'IY INTELLEKT MODELLARINING ISHLASHI VA OPTIMALLASHTIRISH USULLARI

Ruzimboyev Xusniddin Raximberganovich

khusniddindatascientisn1@gmail.com

Avezmatov Ixtiyor Davlatyorovich

Urganch davlat universiteti “Kompyuter ilmlari” kafedrasi.

ixtiyoravezmatov07@gmail.com

https://doi.org/10.5281/zenodo.13924939

Annotatsiya. *Ushbu tezis ma'lumotlar bilan ishlashda sun'iy intellekt (SI) modellarining ishlash printsiplari va ularni optimallashtirish usullarini chuqur o'rganadi. Zamonaviy SI modellarining turli xil ma'lumotlar turlari bilan ishlash qobiliyati, shu jumladan katta hajmli va murakkab strukturali ma'lumotlarni qayta ishlash jarayonlari batafsil tahlil qilinadi. Tezisda modellarning o'qitish algoritmlari, xususan, chuqur o'rganish (deep learning) va mashina o'rganish (machine learning) usullarining qiyosiy tahlili, shuningdek, ularning ishlashini optimallashtirish uchun qo'llaniladigan zamonaviy yondashuvlar va texnikalar keng muhokama qilinadi. Tadqiqot natijalari SI modellarining samaradorligini oshirish va ularning turli sohalardagi amaliy qo'llanilishini kengaytirish imkoniyatlarini ko'rsatadi.*

Kalit so'zlar: *sun'iy intellekt, ma'lumotlarni qayta ishlash, chuqur o'rganish, mashina o'rganish, optimallashtirish usullari, katta ma'lumotlar, gradient tushish, regularizatsiya, transfer o'rganish.*

PERFORMANCE AND OPTIMIZATION METHODS OF ARTIFICIAL INTELLIGENCE MODELS WHEN WORKING WITH DATA

Abstract. *This thesis deeply explores the principles of artificial intelligence (SI) models in data processing and their optimization methods. The ability of modern SI models to work with various types of data, including large-volume and complex structured data processing processes, is analyzed in detail. In the thesis, a comparative analysis of model training algorithms, in particular, deep learning and machine learning methods, as well as modern approaches and techniques used to optimize their performance, is widely discussed. The results of the research show the possibilities of increasing the effectiveness of SI models and expanding their practical application in various fields.*

Keywords: *artificial intelligence, data processing, deep learning, machine learning, optimization methods, big data, gradient descent, regularization, transfer learning.*

ПРОИЗВОДИТЕЛЬНОСТЬ И МЕТОДЫ ОПТИМИЗАЦИИ МОДЕЛЕЙ

ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ПРИ РАБОТЕ С ДАННЫМИ

Аннотация. *В данной диссертации глубоко исследуются принципы работы моделей искусственного интеллекта (СИ) и методы их оптимизации при работе с данными. Подробно анализируется способность современных моделей СИ работать с различными типами данных, включая процессы обработки больших объемов и сложно структурированных данных. В диссертации широко обсуждается сравнительный анализ алгоритмов обучения моделей, в частности методов глубокого обучения и машинного обучения, а также современных подходов и методов, используемых для оптимизации их производительности. Результаты исследования показывают возможности повышения эффективности моделей СИ и расширения их практического применения в различных областях.*

Ключевые слова: *искусственный интеллект, обработка данных, глубокое обучение, машинное обучение, методы оптимизации, большие данные, градиентный спуск, регуляризация, трансферное обучение.*

Kirish

Sun'iy intellekt texnologiyalari kundan-kunga rivojlanib, turli sohalarda keng qo'llanilmoqda. Tibbiyotdan tortib moliyagacha, ta'limdan ishlab chiqarishgacha - SI modellarining qo'llanilish doirasi tobora kengaymoqda. Biroq, SI modellarining samaradorligi ko'p jihatdan ularning ma'lumotlar bilan ishlash qobiliyati va optimallashtirish darajasiga bog'liq.

Katta hajmli va murakkab tuzilishga ega bo'lgan ma'lumotlar bilan ishlash, modellarni tez va samarali o'qitish, shuningdek, ularning aniqligi va ishonchliligini ta'minlash - bularning barchasi zamonaviy SI tadqiqotlarining asosiy yo'nalishlariga aylangan.

Ushbu tadqiqot SI modellarining ishlash printsiplarini chuqur o'rganish va ularni optimallashtirish usullarini har tomonlama tahlil qilish orqali, ushbu texnologiyalarning samaradorligini oshirish yo'llarini aniqlashga qaratilgan. Biz nafaqat mavjud usullarni o'rganamiz, balki yangi yondashuvlarni ham taklif etamiz, bu esa SI modellarining kelajakdagi rivojlanishiga hissa qo'shishi mumkin.

Asosiy qism

SI modellarining ma'lumotlar bilan ishlash printsiplari

1. **Ma'lumotlarni oldindan qayta ishlash (pre-processing):**
   - Ma'lumotlarni tozalash:
     - Yo'qolgan qiymatlarni to'ldirish (imputation)

- Tashqi qiymatlarni (outliers) aniqlash va bartaraf etish
- Noaniq yoki noto'g'ri ma'lumotlarni tuzatish
  - Ma'lumotlarni normallash:
- Min-Max normallash
- Z-score normallash
- Logarifmik transformatsiya
  - Xususiyatlarni ajratish va tanlash:
- Principal Component Analysis (PCA)
- Linear Discriminant Analysis (LDA)
- Recursive Feature Elimination (RFE)
- Genetik algoritmlar orqali xususiyatlarni tanlash

2. **Ma'lumotlarni o'qitish (training):**
   - Nazorat ostida o'qitish (supervised learning):
- Klassifikatsiya algoritmlari (Decision Trees, Random Forest, SVM)
- Regressiya algoritmlari (Linear Regression, Polynomial Regression)
- Ensemble usullar (Bagging, Boosting)
  - Nazoratsiz o'qitish (unsupervised learning):
- Klasterlash algoritmlari (K-Means, DBSCAN, Hierarchical Clustering)
- Dimensiyani kamaytirish usullari (t-SNE, UMAP)
- Associatsiya qoidalari (Apriori algoritmi)
  - Yarim-nazoratli o'qitish (semi-supervised learning):
- Self-training
- Co-training
- Generative models

3. **Model arxitekturalari:**
   - Chuqur neyron tarmoqlari (Deep Neural Networks):
- Fully connected neural networks
- Autoencoders
- Deep Belief Networks (DBN)
  - Konvolyutsion neyron tarmoqlari (Convolutional Neural Networks):
- LeNet, AlexNet, VGGNet
- ResNet, Inception, DenseNet
- U-Net (segmentatsiya uchun)

- Rekurrent neyron tarmoqlari (Recurrent Neural Networks):
  - LSTM (Long Short-Term Memory)
  - GRU (Gated Recurrent Unit)
  - Bidirectional RNN
- Transformerlar:
  - BERT, GPT, T5
  - Vision Transformers (ViT)

Optimallashtirish usullari

1. **Gradient tushish (Gradient Descent) va uning variantlari:**
   - Stoxastik gradient tushish (Stochastic Gradient Descent):
     - Mini-batch SGD
     - Momentum SGD
   - Adaptive o'qitish stavkasi usullari:
     - AdaGrad
     - RMSprop
     - Adam optimizatori
     - AdamW (weight decay bilan)
   - Ikkinchi tartibli usullar:
     - Newton's method
     - Quasi-Newton methods (BFGS, L-BFGS)
2. **Regularizatsiya texnikalari:**
   - L1 va L2 regularizatsiya:
     - Lasso regression (L1)
     - Ridge regression (L2)
     - Elastic Net (L1 + L2)
   - Dropout:
     - Standart dropout
     - Spatial dropout
     - Variational dropout
   - Batch normalization:
     - Layer normalization
     - Instance normalization
     - Group normalization

- Early stopping
- Data augmentation

3. **Giper-parametrlarni sozlash:**

- Grid qidiruv (Grid Search)
- Random qidiruv (Random Search)
- Bayesian optimallashtirish:
  - Gaussian Process-based Bayesian Optimization
  - Tree-structured Parzen Estimators (TPE)
- Evolyutsion algoritmlar:
  - Genetik algoritmlar
  - Particle Swarm Optimization (PSO)

Zamonaviy yondashuvlar

1. **Transfer o'rganish (Transfer Learning):**

- Fine-tuning
- Feature extraction
- Domain adaptation
- Multi-task learning

2. **Meta-o'rganish (Meta-Learning):**

- Model-Agnostic Meta-Learning (MAML)
- Reptile
- Prototypical Networks

3. **Avtomatik mashina o'rganish (AutoML):**

- Neural Architecture Search (NAS)
- Hyperparameter optimization
- Automated feature engineering
- Auto-sklearn, Auto-PyTorch

4. **Federated Learning:**

- Horizontally Federated Learning
- Vertically Federated Learning
- Federated Transfer Learning

5. **Explainable AI (XAI):**

- SHAP (SHapley Additive exPlanations)
- LIME (Local Interpretable Model-agnostic Explanations)

- Integrated Gradients

Kutilayotgan natijalar

Tadqiqot natijasida quyidagi natijalarga erishish kutilmoqda:

1. SI modellarining turli xil ma'lumotlar turlari bilan ishlash samaradorligini baholash va taqqoslash.

2. Optimallashtirish usullarining qiyosiy tahlili va ularning turli vaziyatlarda qo'llanilishi bo'yicha batafsil tavsiyalar ishlab chiqish.

3. Zamonaviy yondashuvlarning SI modellarining ishlashiga ta'sirini o'rganish va ularning amaliy qo'llanilish sohalarini aniqlash.

4. Katta hajmli va murakkab strukturali ma'lumotlar bilan ishlashda SI modellarining samaradorligini oshirish uchun yangi yondashuvlar va algoritmlar taklif etish.

5. O'zbek tilidagi ma'lumotlar bilan ishlash uchun maxsus SI modellarini yaratish va optimallashtirish bo'yicha tavsiyalar ishlab chiqish.

Xulosa

Ma'lumotlar bilan ishlashda sun'iy intellekt modellarining ishlashi va optimallashtirish usullarini chuqur o'rganish orqali, biz ushbu texnologiyalarning samaradorligini sezilarli darajada oshirish imkoniyatiga ega bo'lamiz. Bu esa, o'z navbatida, SI modellarining turli sohalarda, jumladan tibbiyot, moliya, ta'lim va ishlab chiqarish kabi sohalarda yanada keng va samarali qo'llanilishiga olib keladi.

Tadqiqot davomida o'rganilgan va taklif etilgan usullar nafaqat mavjud muammolarni hal qilishga yordam beradi, balki SI texnologiyalarining yangi imkoniyatlarini ham ochib beradi.

Masalan, transfer o'rganish va meta-o'rganish usullari kam ma'lumotli vaziyatlarda ham samarali modellar yaratish imkonini bersa, federated learning ma'lumotlar xavfsizligi va maxfiyligini ta'minlagan holda turli manbalardan o'rganish imkoniyatini yaratadi.

Kelajakda, yangi optimallashtirish usullari va ma'lumotlarni qayta ishlash texnikalarining rivojlanishi bilan SI modellarining imkoniyatlari yanada kengayishi kutilmoqda. Bu esa, o'z navbatida, yangi ilmiy kashfiyotlar, innovatsion mahsulotlar va xizmatlarning paydo bo'lishiga olib keladi. Shu bilan birga, SI modellarining shaffofligini ta'minlash va ularning qarorlarini tushuntirish imkoniyatini beruvchi Explainable AI kabi yo'nalishlar ham tobora muhim ahamiyat kasb etmoqda.

Xulosa qilib aytganda, ushbu tadqiqot nafaqat SI modellarining texnik jihatlarini yaxshilashga, balki ularning jamiyat hayotidagi rolini oshirishga ham xizmat qiladi. Bu esa, o'z

navbatida, insoniyat oldida turgan global muammolarni hal qilishda SI texnologiyalaridan yanada samarali foydalanish imkoniyatini yaratadi.

REFERENCES

1. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.
2. Ruder, S. (2016). An overview of gradient descent optimization algorithms. arXiv preprint arXiv:1609.04747.
3. Hutter, F., Kotthoff, L., & Vanschoren, J. (2019). Automated machine learning: methods, systems, challenges. Springer Nature.
4. Wang, T., Zhu, J. Y., Torralba, A., & Efros, A. A. (2018). Dataset distillation. arXiv preprint arXiv:1811.10959.
5. Shalev-Shwartz, S., & Ben-David, S. (2014). Understanding machine learning: From theory to algorithms. Cambridge university press.
6. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.
7. Finn, C., Abbeel, P., & Levine, S. (2017). Model-agnostic meta-learning for fast adaptation of deep networks. In International Conference on Machine Learning (pp. 1126-1135). PMLR.
8. Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. ACM Transactions on Intelligent Systems and Technology (TIST), 10(2), 1-19.