

ZAMONAVIY SUN'IY INTELLEKT ALGORITMLARI VA ULARNING ARXITEKTURAVIY TAHLILI

¹Primbetov Abbaz, ²Primbetov Aziz

¹PhD talaba, Muhammad al Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti

¹Katta o'qituvchi, Toshkent amaliy fanlar universiteti

²Nukus davlat texnika universiteti

abbaz0203@mail.ru, aziz2212@mail.ru

Annotatsiya - Ushbu maqolada zamonaviy sun'iy intellekt algoritmlarining nazariy asoslari, matematik modellari va amaliy qo'llanilish sohalari tahlil qilinadi. Tadqiqotda klassik mashinaviy o'rganish algoritmlari, jumladan Random Forest, Support Vector Machine (SVM) va Gradient Boosting/XGBoost modellarining ishlash prinsiplari ko'rib chiqilgan. Shuningdek, chuqur o'rganish (Deep Learning) arxitekturalari — CNN, RNN, LSTM, Transformer va Graph Neural Network (GNN) modellarining tuzilishi, afzalliklari va cheklovlari ilmiy jihatdan yoritilgan. Tadqiqot natijalari sun'iy intellekt algoritmlarining robototexnika, signal va tasvirlarga ishlov berish, aqlli materiallar hamda Big Data texnologiyalaridagi yuqori samaradorligini ko'rsatadi. Bundan tashqari, Explainable AI (XAI) va Transfer Learning yondashuvlarining zamonaviy AI tizimlaridagi ahamiyati tahlil qilingan.

Kalit so'zlar: Deep Learning, CNN, RNN, LSTM, Transformer.

Kirish

Klassik modellarning yuqorida keltirilgan cheklovlari (high-dimensional data bilan ishlashdagi qiyinchiliklar) neyron tarmoqlari ierarxiyasiga asoslangan chuqur o'rganish usullariga o'tishni taqozo etadi.

Chuqur o'rganish (Deep Learning) algoritmlari zamonaviy sun'iy intellektning eng muhim yo'nalishlaridan biri hisoblanadi. Ushbu algoritmlar ko'p qatlamli neyron tarmoqlar asosida ishlaydi va katta hajmdagi ma'lumotlardan murakkab hamda yashirin bog'lanishlarni avtomatik ravishda aniqlash imkonini beradi. An'anaviy mashinaviy o'rganish usullaridan farqli ravishda, Deep Learning modellarida xususiyatlarni (feature extraction) qo'lda ajratish talab qilinmaydi, chunki modelning o'zi ma'lumotlarning ierarxik reprezentatsiyasini shakllantiradi. Chuqur o'rganish modellarining umumiy matematik modeli quyidagicha ifodalanadi:

$$a^{(l)} = \sigma(W^{(l)}a^{(l-1)} + b^{(l)})$$

- — l -qatlam chiqishi;
- $W^{(l)}$ — vaznlar matritsasi;
- $b^{(l)}$ — siljish vektori;
- σ — aktivatsiya funksiyasi.

Modelni o'qitish jarayoni quyidagi yo'qotish funksiyasini minimallashtirishga asoslanadi:

$$L(\theta) = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

Parametrlarni yangilash gradient tushishi algoritmi yordamida amalga oshiriladi:

$$W = W - \eta \frac{\partial L}{\partial W}$$

Deep Learning arxitekturalari

Deep Learning arxitekturalari ma'lumot turi va masala xarakteriga qarab turli modellar asosida quriladi.

1. Konvolyutsion neyron tarmoqlari (CNN).

Konvolyutsion neyron tarmoqlari (Convolutional Neural Networks — CNN) fazoviy bog'lanishlarni aniqlashda eng samarali yondashuvlardan biri hisoblanadi. Ushbu modellar ayniqsa tasvirlarga ishlov berish va materiallarning mikroskopik strukturasini tahlil qilishda keng qo'llaniladi.

CNN arxitekturasining asosiy elementi konvolyutsiya operatsiyasi hisoblanadi:

$$h_{i,j} = \sum_{m,n} X_{i+m,j+n} K_{m,n}$$

bu yerda:

- X — kirish tasviri;
- K — filtr (kernel);
- $h_{i,j}$ — feature map elementi.

Konvolyutsiya jarayonida filtr tasvir bo'ylab harakatlanib, lokal fazoviy xususiyatlarni aniqlaydi. Ushbu yondashuv material yuzasidagi mikroyoriqlar, deformatsiyalar va struktura nuqsonlarini aniqlashda samarali hisoblanadi.

CNN modellarida aktivatsiya funksiyasi sifatida ko'pincha ReLU funksiyasi qo'llaniladi:

$$f(x) = \max(0, x)$$

Pooling operatsiyasi ma'lumot o'lchamini kamaytirish va asosiy xususiyatlarni saqlab qolish uchun ishlatiladi:

$$p_{i,j} = \max_{(m,n) \in R} h_{m,n}$$

CNN algoritmlari tasviriy ma'lumotlarni tahlil qilishda keng qo'llaniladi. Ushbu modellar material yuzasidagi nuqsonlarni aniqlash (material defect detection), kristall strukturalarni klassifikatsiya qilish, SEM/TEM mikroskopik tasvirlarini qayta ishlash hamda mikroyoriqlar va fazoviy segmentatsiyani aniqlashda yuqori samaradorlikka ega. Ayniqsa, konvolyutsion qatlamlar yordamida obyektlarning lokal fazoviy xususiyatlari avtomatik ravishda ajratib olinadi, bu esa murakkab strukturalarni yuqori aniqlik bilan tahlil qilish imkonini beradi.

CNN modellarining asosiy afzalligi tasvirlardan avtomatik feature extraction bajarishi va fazoviy bog'lanishlarni chuqur o'rganish qobiliyatidir. Shu sababli ular aqlli materiallar va intellektual tizimlarning mikrostrukturaviy tasvirlarini tahlil qilishda istiqbolli yondashuv hisoblanadi. Biroq ushbu modellar katta hajmdagi o'quv ma'lumotlari hamda yuqori hisoblash resurslarini talab qiladi (1-rasm).

2. Rekurrent neyron tarmoqlari (RNN) va LSTM.

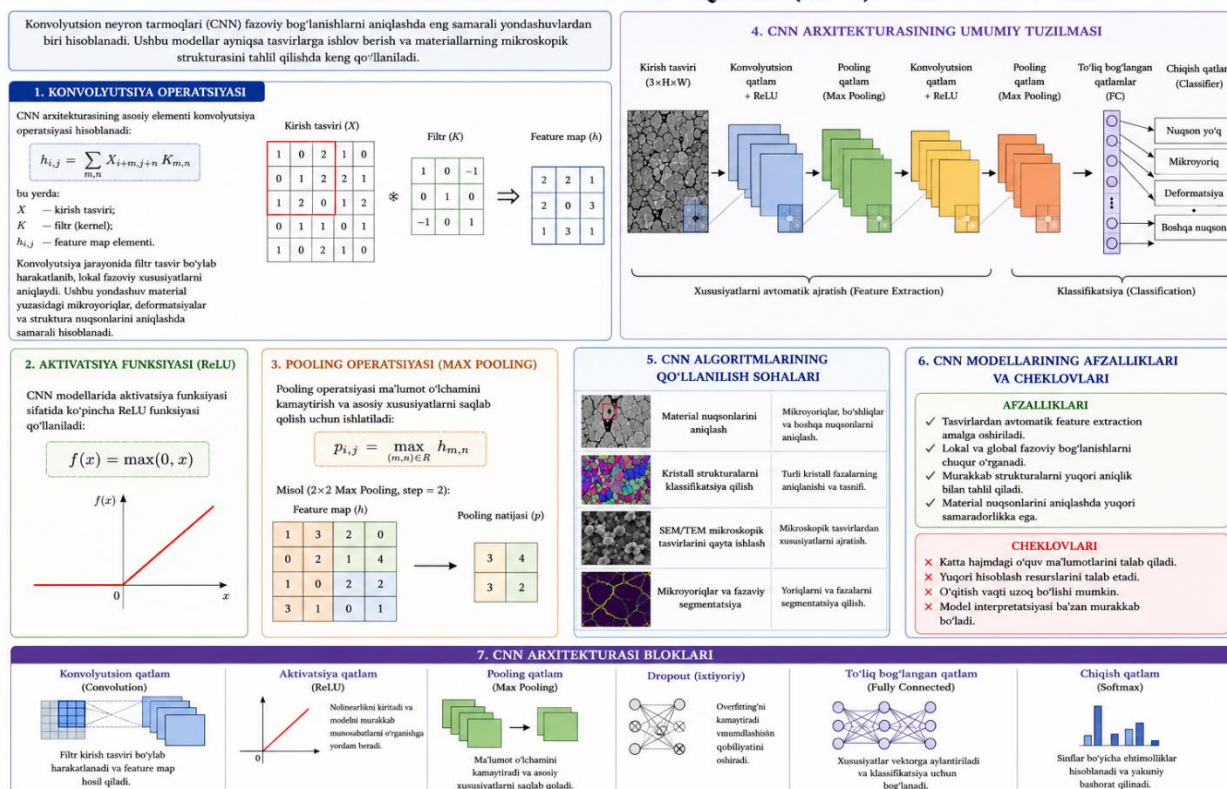
Rekurrent neyron tarmoqlari (Recurrent Neural Networks — RNN) vaqtga bog'liq ma'lumotlarni qayta ishlash uchun mo'ljallangan. Ushbu modellar oldingi holat haqidagi ma'lumotni saqlab qolish imkoniga ega bo'lib, vaqt qatorlari va dinamik signallarni tahlil qilishda samarali hisoblanadi. RNN modelining umumiy matematik modeli:

$$h_t = f(W_h h_{t-1} + W_x x_t + b)$$

bu yerda:

- x_t — joriy kirish signali;
- h_t — yashirin holat;
- W_h, W_x — vazn matritsalar;
- b — bias parametri.

KONVOLYUTSION NEYRON TARMOQLARI (CNN) ARXITEKTURASI



1-rasm: Konvolyutsion neyron tarmoqlari (CNN) arxitekturasini

RNN modellarida uzoq ketma-ketliklarda gradientning yo'qolishi (vanishing gradient) muammosi kuzatiladi. Ushbu muammoni hal qilish uchun Long Short-Term Memory (LSTM) arxitekturasini ishlab chiqilgan. LSTM ning forget gate modeli:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$

Cell state yangilanishi:

$$C_t = f_t C_{t-1} + i_t \tilde{C}_t$$

LSTM modellarining asosiy afzalligi uzoq muddatli bog'lanishlarni saqlab qolish va vaqt bo'yicha o'zgaruvchi signallar orasidagi yashirin bog'liqliklarni aniqlash imkoniyatiga egaligidir. Ushbu modellar vaqt qatorlari ko'rinishidagi ma'lumotlarni tahlil qilishda yuqori samaradorlikni namoyon qiladi hamda gradientning yo'qolishi muammosini kamaytiradi. Shu sababli LSTM arxitekturasini piezo-sensor signallarini qayta ishlash, vibratsiya monitoringi, predictive maintenance tizimlari, dinamik deformatsiyalarni bashoratlash va aqlli aktuator tizimlari monitoringida keng qo'llaniladi.

RNN va LSTM modellar vaqtga bog'liq jarayonlarni modellashtirishda samarali hisoblanadi, biroq katta uzunlikdagi ketma-ketliklarda hisoblash murakkabligi ortishi va o'qitish jarayonining sekinlashishi kuzatiladi (2-rasm).

3. Transformer arxitekturasini.

Transformer modellar zamonaviy Deep Learning arxitekturalarining eng istiqbolli yo'nalishlaridan biri hisoblanadi. Ushbu modellar self-attention mexanizmi asosida ishlaydi va uzoq masofali bog'lanishlarni aniqlashda yuqori samaradorlikka ega. Self-attention mexanizmi:

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

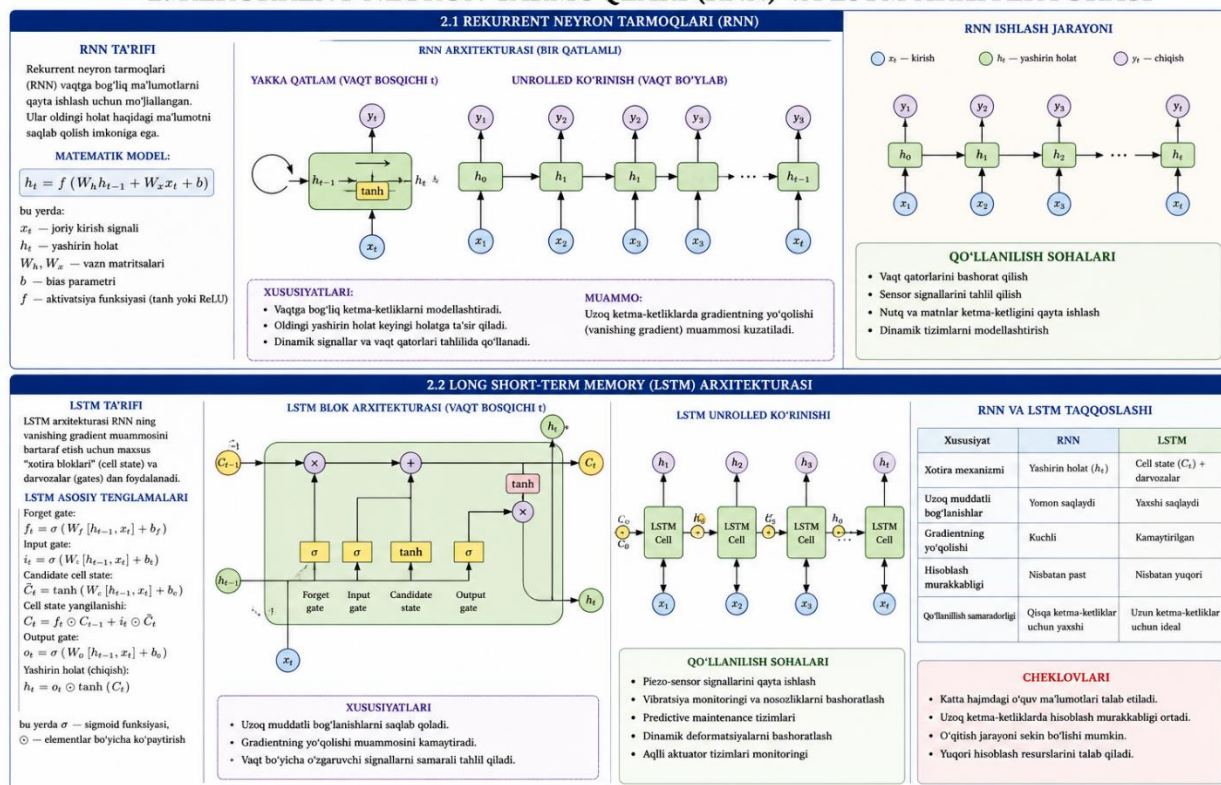
Transformer modellarida Multi-head Attention mexanizmi parallel ravishda bir nechta bog'lanishlarni o'rganish imkonini beradi:

$$MultiHead(Q, K, V) = Concat(head_1, \dots, head_h)W^O$$

Pozitsion kodlash modeli:

$$PE(pos, 2i) = \sin\left(\frac{pos}{10000^{2i/d}}\right)$$

2. REKURRENT NEYRON TARMOQLARI (RNN) VA LSTM ARXITEKTURASI



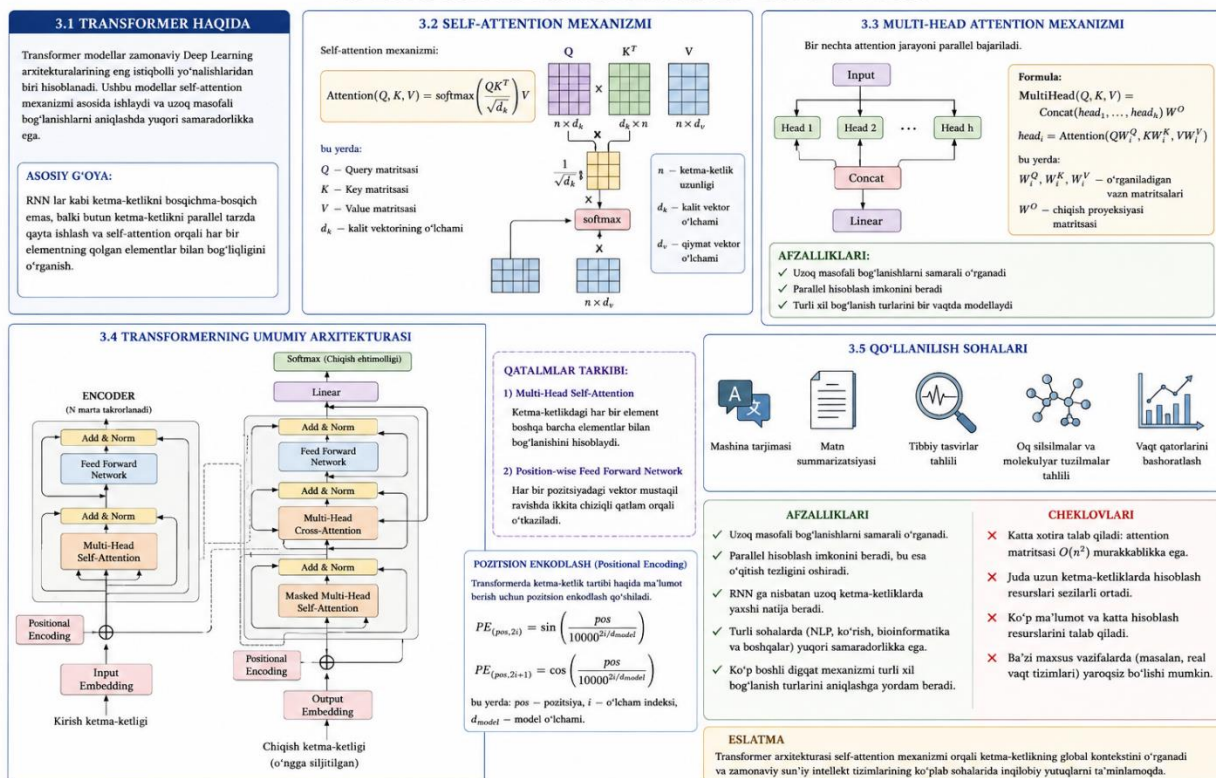
2- rasm. RNN va LSTM arxitekturas

Transformer modellarning asosiy afzalligi uzoq masofali bog'lanishlarni samarali aniqlash hamda ma'lumotlarni parallel qayta ishlash imkoniyatiga egaligidir. Ushbu modellar multimodal ma'lumotlar va katta hajmdagi datasetlar bilan ishlashda yuqori samaradorlikni namoyon qiladi. Shu sababli ular sensor signallarini tahlil qilish, murakkab fizik jarayonlarni bashoratlash va material xossalarini prognozlash masalalarida keng qo'llaniladi.

Biroq Transformer modellarning asosiy kamchiligi katta hisoblash resurslari va yuqori hajmdagi o'quv ma'lumotlarini talab qilishidir.

Model	Afzalligi	Kamchiligi	Qo'llanilish sohasi
CNN	Fazoviy belgilarni aniqlashda yuqori aniqlik	Katta hisoblash resurslari talab qiladi	Tasvirlar va mikroyoriqlar tahlili
RNN	Vaqt qatorlarini qayta ishlash	Gradient yo'qolishi muammosi	Sensor signallari
LSTM	Uzoq bog'lanishlarni saqlaydi	Hisoblash murakkabligi yuqori	Predictive maintenance
Transformer	Long-range dependency tahlili	Juda katta dataset talab qiladi	Multimodal va fizik signallar

3. TRANSFORMER ARXITEKTURASI



3-rasm- Transformer arxitekturasida

1-jadval. Chuqur o'rganish arxitekturalarining qiyosiy tahlili

Zamonaviy sun'iy intellekt algoritmlari ma'lumot turi va qo'llanilish sohasi bo'yicha turlicha samaradorlikka ega. Klassik mashinaviy o'rganish algoritmlari, jumladan XGBoost va Random Forest modellar kichik va o'rta hajmdagi ma'lumotlarda yuqori aniqlik hamda tezkor ishlash imkonini beradi. Biroq murakkab nolinear va yuqori o'lchamli ma'lumotlarda ularning samaradorligi cheklanishi mumkin.

Xulosa

Mazkur tadqiqotda zamonaviy sun'iy intellekt algoritmlarining nazariy asoslari, matematik modellari va arxitekturalari xususiyatlari tahlil qilindi. Tadqiqot davomida klassik mashinaviy o'rganish algoritmlari bilan bir qatorda CNN, RNN, LSTM va Transformer kabi chuqur o'rganish arxitekturalarining ishlash prinsiplari, afzalliklari va cheklovlari o'rganildi. Natijalar shuni ko'rsatdiki, CNN modellar tasviriy ma'lumotlarni tahlil qilishda, RNN va LSTM algoritmlari vaqt qatorlari hamda sensor signallarini qayta ishlashda, Transformer arxitekturalari esa katta hajmdagi va multimodal ma'lumotlarni tahlil qilishda yuqori samaradorlikka ega.

Shuningdek, zamonaviy AI algoritmlarining aqlli materiallar, intellektual tizimlar, robototexnika va Big Data texnologiyalaridagi amaliy ahamiyati yoritildi. Tahlillar sun'iy intellekt modellarining murakkab bog'lanishlarni aniqlash va bashoratlash imkoniyatlari yuqori ekanligini ko'rsatdi. Biroq ushbu modellar katta hajmdagi ma'lumotlar va yuqori hisoblash resurslarini talab qilishi ularning asosiy cheklovlaridan biri hisoblanadi.

Kelgusida Explainable AI, Transfer Learning va resurs tejamkor sun'iy intellekt modellarini rivojlantirish zamonaviy ilmiy tadqiqotlarning muhim yo'nalishlaridan biri bo'lib qoladi. Bu esa sun'iy intellekt tizimlarining aniqligi, ishonchliligi va real vaqt tizimlaridagi samaradorligini yanada oshirishga xizmat qiladi.

Foydalanilgan adabiyotlar

- [1] Rogers, C. A. (1989). Intelligent materials. *Scientific American*, 261(3), 154–157.
- [2] Gandhi, M. V., & Thompson, B. S. (1992). *Smart Materials and Structures*. Chapman & Hall.
- [3] Srinivasan, A. V., & McFarland, D. M. (2001). *Smart Structures: Analysis and Design*. Cambridge University Press.
- [4] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- [5] Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
- [6] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- [7] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297.
- [8] Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189–1232.
- [9] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). ACM.
- [10] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444.
- [11] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780.
- [12] Vaswani, A., Shazeer, N., Parmar, N., et al. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems* (pp. 5998–6008).
- [13] Kipf, T. N., & Welling, M. (2017). Semi-supervised classification with graph convolutional networks. *International Conference on Learning Representations (ICLR)*.
- [14] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems* (pp. 4765–4774).
- [15] Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359.
- [16] Dean, J., & Ghemawat, S. (2008). MapReduce: Simplified data processing on large clusters. *Communications of the ACM*, 51(1), 107–113.
- [17] Zaharia, M., Chowdhury, M., Franklin, M. J., Shenker, S., & Stoica, I. (2010). Spark: Cluster computing with working sets. In *Proceedings of the 2nd USENIX Conference on Hot Topics in Cloud Computing*.
- [18] Kreps, J., Narkhede, N., & Rao, J. (2011). Kafka: A distributed messaging system for log processing. In *Proceedings of the NetDB Workshop*.