

CK0255/CKP8377 - Aprendizado Profundo

Proposta para o Projeto Final

Prof. Eliezer de Souza da Silva

Semestre: 2025.2

Visão Geral e Objetivos

O projeto final desta disciplina será um **seminário de pesquisa combinado com uma implementação prática**, no estilo de um ciclo de pesquisa acadêmica. O objetivo é que cada grupo realize um mergulho profundo em um artigo de pesquisa influente da área de Aprendizado Profundo, analisando-o criticamente e, crucialmente, implementando sua contribuição central.

Os principais objetivos desta atividade são:

- Desenvolver a habilidade de ler, interpretar e entender artigos técnicos de pesquisa.
- Ganhar experiência prática na implementação de modelos e algoritmos a partir de publicações científicas.
- Aprender a analisar criticamente um trabalho científico, identificando suas contribuições, pontos fortes e fraquezas.
- Praticar a comunicação de ideias técnicas complexas de forma clara e concisa.

A Estrutura do Projeto

O projeto será desenvolvido em grupos de 3 pessoas e seguirá os seguintes passos:

Passo 1: Seleção e Leitura Cada grupo deve explorar e ler **3 a 4 artigos**. Você pode incluir artigos de fora da lista sugerida, mas precisa incluir pelo menos um da lista sugerida.

Passo 2: Escolha e Proposta O grupo escolhe **um único artigo principal** e submete uma breve proposta de 2 páginas contendo:

- O artigo escolhido e os integrantes do grupo.
- Um resumo da principal contribuição do artigo.
- Um plano para a implementação: Qual é a parte central do artigo que vocês irão implementar? Qual dataset (pode ser um "toy dataset" como MNIST/CIFAR-10 ou um mais simples) vocês usarão para validar a implementação?
- Um resumo dos outros artigos escolhidos pelo grupo, como uma mini revisão de literatura.

Passo 3: Implementação e Análise Experimental O grupo deve implementar a contribuição chave do artigo. O objetivo é criar uma prova de conceito funcional que demonstre o entendimento da ideia central.

Passo 4: Preparação e Apresentação O grupo prepara uma apresentação de **20-25 minutos** e a apresenta no seminário final. Um relatório final e o código-fonte final (Jupyter Notebook) deve ser submetido juntamente com os slides.

O Conteúdo da Apresentação

A apresentação deve ser didática, assumindo que a audiência conhece os fundamentos de Deep Learning, mas não leu o artigo. O conteúdo deve incluir:

1. **Contextualização:** O problema que o artigo aborda e sua importância.
2. **Contribuição Principal:** A ideia central, método ou arquitetura proposta.
3. **Implementação:** Uma breve visão geral da sua implementação e os desafios encontrados.
4. **Resultados:** Demonstração dos resultados da sua implementação (gráficos, tabelas, visualizações).
5. **Análise Crítica:** Pontos fortes, fracos e limitações do artigo.
6. **Extensões:** Propostas para trabalhos futuros ou soluções alternativas.

Critérios de Avaliação

O projeto final será avaliado com base em dois componentes:

- **Apresentação (50%):**
 - Clareza da explicação da contribuição do artigo.
 - Profundidade da análise crítica.
 - Qualidade da apresentação e das respostas às perguntas.
- **Relatório, Implementação e Notebook (50%):**
 - Correção e funcionalidade do código implementado.
 - Clareza, organização e reproduzibilidade.
 - Análise e discussão dos resultados experimentais obtidos.

Lista de Artigos Sugeridos

A lista é organizada por categoria. O grupo deve citar e discutir o trabalho em sua apresentação.

Categoria 1: Fundamentos e Teoria

1. Montufar et al. (2014). *On the number of linear regions of deep neural networks.* (Montúfar et al. 2014)
2. Belkin et al. (2019). *Reconciling modern machine-learning practice and the classical bias-variance trade-off.* (Belkin et al. 2019)
3. Jacot et al. (2018). *Neural Tangent Kernel: Convergence and Generalization in Neural Networks.* (Jacot, Gabriel e Hongler 2018)

Categoria 2: Arquiteturas

4. Krizhevsky, Sutskever & Hinton (2012). *ImageNet Classification with Deep Convolutional Neural Networks (AlexNet).* (Krizhevsky, Sutskever e Hinton 2012)
5. Simonyan & Zisserman (2014). *Very Deep Convolutional Networks for Large-Scale Image Recognition (VGG).* (Simonyan e Zisserman 2014)
6. He et al. (2016). *Deep Residual Learning for Image Recognition (ResNet).* (He et al. 2016)
7. Hochreiter & Schmidhuber (1997). *Long short-term memory (LSTM).* (Hochreiter e Schmidhuber 1997b)
8. Vaswani et al. (2017). *Attention Is All You Need (Transformer).* (Vaswani et al. 2017)

Categoria 3: Modelos Generativos

9. Goodfellow et al. (2014). *Generative Adversarial Networks (GANs).* (Goodfellow et al. 2014)
10. Kingma & Welling (2013). *Auto-Encoding Variational Bayes (VAE).* (Kingma e Welling 2013)
11. Karras et al. (2019). *A Style-Based Generator Architecture for Generative Adversarial Networks (StyleGAN).* (Karras, Laine e Aila 2019)
12. Ho, Jain & Abbeel (2020). *Denoising Diffusion Probabilistic Models (DDPM).* (Ho, Jain e Abbeel 2020)

Categoria 4: Otimização e Regularização

13. Kingma & Ba (2014). *Adam: A Method for Stochastic Optimization.* (Kingma e Ba 2015)
14. Srivastava et al. (2014). *Dropout: A simple way to prevent neural networks from overfitting.* (Srivastava et al. 2014)
15. Ioffe & Szegedy (2015). *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.* (Ioffe e Szegedy 2015)
16. Hochreiter & Schmidhuber (1997). *Flat Minima.* (Hochreiter e Schmidhuber 1997a)

Categoría 5: Grafos e Sistemas Físicos

17. Kipf & Welling (2016). *Semi-Supervised Classification with Graph Convolutional Networks (GCN)*. (Kipf e Welling 2016)
18. Wu et al. (2019). *Simplifying Graph Convolutional Networks (SGC)*. (Wu et al. 2019)
19. Mesquita et al. (2020). *Rethinking pooling in graph neural networks*. (Mesquita, Souza e Kaski 2020)
20. Raissi et al. (2019). *Physics-Informed Neural Networks (PINNs)*. (Raissi, Perdikaris e Karniadakis 2019)
21. Chen et al. (2018). *Neural Ordinary Differential Equations (Neural ODEs)*. (R. T. Chen et al. 2018)

Categoría 6: Paradigmas de Aprendizagem

22. Silver et al. (2016). *Mastering the game of Go with deep neural networks and tree search (AlphaGo)*. (Silver et al. 2016)
23. Devlin et al. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. (Devlin et al. 2019)
24. Brown et al. (2020). *Language Models are Few-Shot Learners (GPT-3)*. (Brown et al. 2020)
25. Chen et al. (2020). *A Simple Framework for Contrastive Learning of Visual Representations (SimCLR)*. (T. Chen et al. 2020)
26. Radford et al. (2021). *Learning Transferable Visual Models From Natural Language Supervision (CLIP)*. (Radford et al. 2021)
27. Jumper et al. (2021). *Highly accurate protein structure prediction with AlphaFold*. (Jumper et al. 2021)
28. Mikolov et al. (2013). *Efficient Estimation of Word Representations in Vector Space (Word2Vec)*. (Mikolov et al. 2013)

Referências

- Belkin, Mikhail et al. (2019). “Reconciling modern machine-learning practice and the classical bias-variance trade-off”. Em: *Proceedings of the National Academy of Sciences* 116.32, pp. 15849–15854.
- Brown, Tom B et al. (2020). “Language models are few-shot learners”. Em: *arXiv preprint arXiv:2005.14165*.
- Chen, Ricky TQ et al. (2018). “Neural ordinary differential equations”. Em: *Advances in neural information processing systems*, pp. 6571–6583.
- Chen, Ting et al. (2020). “A simple framework for contrastive learning of visual representations”. Em: *International conference on machine learning*, pp. 900–900.
- Devlin, Jacob et al. (2019). “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”. Em: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pp. 4171–4186.
- Goodfellow, Ian et al. (2014). “Generative adversarial nets”. Em: *Advances in neural information processing systems*, pp. 2672–2680.

- He, Kaiming et al. (2016). “Deep residual learning for image recognition”. Em: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778.
- Ho, Jonathan, Ajay Jain e Pieter Abbeel (2020). “Denoising diffusion probabilistic models”. Em: *Advances in neural information processing systems*. Vol. 33, pp. 684–695.
- Hochreiter, Sepp e Jürgen Schmidhuber (1997a). “Flat maxima”. Em: *Neural computation* 9.1, pp. 1–42.
- (1997b). “Long short-term memory”. Em: *Neural computation* 9.8, pp. 1735–1780.
- Ioffe, Sergey e Christian Szegedy (2015). “Batch normalization: Accelerating deep network training by reducing internal covariate shift”. Em: *arXiv preprint arXiv:1502.03167*.
- Jacot, Arthur, Franck Gabriel e Clément Hongler (2018). “Neural Tangent Kernel: Convergence and Generalization in Neural Networks”. Em: *Advances in Neural Information Processing Systems*. Vol. 31.
- Jumper, John et al. (2021). “Highly accurate protein structure prediction with AlphaFold”. Em: *Nature* 596.7873, pp. 583–589.
- Karras, Tero, Samuli Laine e Timo Aila (2019). “A style-based generator architecture for generative adversarial networks”. Em: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 4401–4410.
- Kingma, Diederik P e Jimmy Ba (2015). “Adam: A method for stochastic optimization”. Em: *International Conference on Learning Representations*.
- Kingma, Diederik P e Max Welling (2013). “Auto-encoding variational bayes”. Em: *arXiv preprint arXiv:1312.6114*.
- Kipf, Thomas N e Max Welling (2016). “Semi-supervised classification with graph convolutional networks”. Em: *arXiv preprint arXiv:1609.2907*.
- Krizhevsky, Alex, Ilya Sutskever e Geoffrey E Hinton (2012). “Imagenet classification with deep convolutional neural networks”. Em: *Advances in neural information processing systems*, pp. 1097–1105.
- Mesquita, Diego, Amauri Souza e Samuel Kaski (2020). “Rethinking pooling in graph neural networks”. Em: *Advances in Neural Information Processing Systems* 33, pp. 2220–2231.
- Mikolov, Tomas et al. (2013). “Efficient estimation of word representations in vector space”. Em: *Workshop at ICLR*.
- Montúfar, Guido F et al. (2014). “On the number of linear regions of deep neural networks”. Em: *Advances in neural information processing systems* 27.
- Radford, Alec et al. (2021). “Learning transferable visual models from natural language supervision”. Em: *International Conference on Machine Learning*.
- Raiissi, Maziar, Paris Perdikaris e George E Karniadakis (2019). “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations”. Em: *Journal of Computational Physics*. Vol. 378, pp. 686–707.
- Silver, David et al. (2016). “Mastering the game of Go with deep neural networks and tree search”. Em: *Nature* 529.7587, pp. 484–489.
- Simonyan, Karen e Andrew Zisserman (2014). “Very deep convolutional networks for large-scale image recognition”. Em: *arXiv preprint arXiv:1409.1556*.
- Srivastava, Nitish et al. (2014). “Dropout: A simple way to prevent neural networks from overfitting”. Em: *The Journal of Machine Learning Research* 15.1, pp. 1929–1958.
- Vaswani, Ashish et al. (2017). “Attention Is All You Need”. Em: *Advances in neural information processing systems*, pp. 5998–6008.
- Wu, Felix et al. (2019). “Simplifying Graph Convolutional Networks”. Em: *International Conference on Machine Learning (ICML)*.