The Ultimate Must-Read AI Research Paper List

For Serious AI/ML/DL/NLP/CV/GAN/LLM Practitioners

ABSOLUTE MUST-KNOWS *(Start Here - These 15 Papers Define Modern AI)*

1. Attention Is All You Need | Vaswani et al. | 2017 [1]

- **Core Foundation:** Introduced Transformer architecture, powering all modern LLMs
- **Citations:** 173,000+ | **Impact:** Revolutionary - enabled ChatGPT, GPT-4, BERT
- **Link:** arXiv | Code

2. ImageNet Classification with Deep Convolutional Neural Networks (AlexNet) | Krizhevsky, Sutskever, Hinton | 2012 [1]

- **Core Foundation:** Started the modern deep learning revolution
- **Citations:** 100,000+ | **Impact:** Made deep learning famous, sparked AI boom
- **Link:** Paper | Code

3. Generative Adversarial Networks | Ian Goodfellow et al. | 2014 [1]

- **Core Foundation:** Introduced GAN framework for generative modeling
- **Citations:** 60,000+ | **Impact:** Revolutionized image generation, synthetic data
- **Link:** arXiv | Code

4. Deep Residual Learning for Image Recognition (ResNet) | He et al. | 2015 [1]

- **Core Foundation:** Solved vanishing gradient problem with residual connections
- **Citations:** 200,000+ | **Impact:** Enabled training very deep networks
- **Link:** arXiv | Code

5. BERT: Pre-training of Deep Bidirectional Transformers | Devlin et al. | 2018 [1]

- **Core Foundation:** Bidirectional language understanding, pre-training paradigm
- **Citations:** 80,000+ | **Impact:** Transformed NLP, established transfer learning
- **Link:** arXiv | Code

6. Human-level control through deep reinforcement learning (DQN) | Mnih et al. | 2015 [2]

- **Core Foundation:** Combined deep learning with reinforcement learning

- **Citations:** 37,000+ | **Impact:** Achieved human-level game performance
- **Link:** Nature | Code

7. Adam: A Method for Stochastic Optimization | Kingma & Ba | 2014 [3]

- **Core Foundation:** Most widely used optimization algorithm in deep learning
- **Citations:** 226,000+ | **Impact:** Default optimizer for neural networks
- **Link:** arXiv | Code

8. Playing Atari with Deep Reinforcement Learning | Mnih et al. | 2013 [4]

- **Core Foundation:** First successful deep RL, introduced DQN concept
- **Citations:** 15,000+ | **Impact:** Proved RL could work with raw pixels
- **Link:** arXiv | Code

9. Neural Machine Translation by Jointly Learning to Align and Translate | Bahdanau et al. | 2014 [1]

- **Core Foundation:** Introduced attention mechanism (pre-Transformer)
- **Citations:** 25,000+ | **Impact:** Enabled modern NMT and attention-based models
- **Link:** arXiv | Code

10. Very Deep Convolutional Networks for Large-Scale Image Recognition (VGG) | Simonyan & Zisserman | 2014 [5]

- **Core Foundation:** Established deeper networks with small filters
- **Citations:** 80,000+ | **Impact:** Standard architecture for many CV applications
- **Link:** arXiv | Code

11. Long Short-Term Memory | Hochreiter & Schmidhuber | 1997 [6]

- **Core Foundation:** Solved vanishing gradients in RNNs, enabled sequence modeling
- **Citations:** 100,000+ | **Impact:** Foundation for pre-Transformer NLP
- **Link:** Paper | Code

12. Learning representations by back-propagating errors | Rumelhart, Hinton, Williams | 1986 [6]

- **Core Foundation:** Backpropagation algorithm that enables neural network training
- **Citations:** 80,000+ | **Impact:** Made modern neural networks possible
- **Link:** Paper | Essential theory

13. Auto-Encoding Variational Bayes (VAE) | Kingma & Welling | 2013 [1]

- **Core Foundation:** Principled approach to generative modeling with latent variables
- **Citations:** 30,000+ | **Impact:** Foundation for many generative models
- **Link:** arXiv | Code

14. Language Models are Few-Shot Learners (GPT-3) | Brown et al. | 2020 [7]

- **Core Foundation:** Demonstrated emergent abilities through scaling
- **Citations:** 25,000+ | **Impact:** Sparked current LLM revolution
- **Link:** arXiv | API

15. Scaling Laws for Neural Language Models | Kaplan et al. | 2020[7]

- **Core Foundation:** Mathematical relationship between scale and performance
- **Citations:** 5,000+ | **Impact:** Guides modern LLM development
- **Link:** arXiv | Theory foundation

FOUNDATIONAL THEORY *(Essential Mathematical & Theoretical Foundations)*

16. A Logical Calculus of the Ideas Immanent in Nervous Activity | McCulloch & Pitts | 1943[8]

- **Historical Foundation:** First mathematical model of neural networks
- **Impact:** Established neural computation as viable concept
- **Link:** Paper

17. The Perceptron: A Probabilistic Model | Rosenblatt | 1958[8]

- **Historical Foundation:** First learning algorithm for neural networks
- **Impact:** Established concept of machine learning from data
- **Link:** Paper

18. Computing Machinery and Intelligence | Turing | 1950[1]

- **Historical Foundation:** The Turing Test, philosophical foundations of AI
- **Impact:** Defined what it means for machines to think
- **Link:** Paper

19. Support Vector Networks | Cortes & Vapnik | 1995[1]

- **Core Theory:** Foundation of kernel methods and modern ML theory
- **Citations:** 80,000+ | **Impact:** Pre-deep learning ML workhorse
- **Link:** Paper

20. Random Forests | Breiman | 2001[9]

- **Core Theory:** Ensemble methods and feature importance
- **Citations:** 100,000+ | **Impact:** Still widely used for tabular data
- **Link:** Paper

🧠 DEEP LEARNING FOUNDATIONS *(Papers That Enabled Deep Learning)*

21. Deep Learning | LeCun, Bengio, Hinton | 2015[1]

- **Review Paper:** Comprehensive overview by the "Godfathers of AI"
- **Citations:** 50,000+ | **Impact:** Definitive introduction to deep learning
- **Link:** Nature

22. Gradient-Based Learning Applied to Document Recognition (LeNet) | LeCun et al. | 1998[1]

- **Foundation:** First successful CNN, introduced convolutional layers
- **Citations:** 35,000+ | **Impact:** Blueprint for modern CNNs
- **Link:** Paper

23. A Fast Learning Algorithm for Deep Belief Nets | Hinton et al. | 2006[1]

- **Renaissance:** Revived deep learning after AI winter
- **Citations:** 25,000+ | **Impact:** Started modern deep learning era
- **Link:** Paper

24. Reducing the Dimensionality of Data with Neural Networks | Hinton & Salakhutdinov | 2006[1]

- **Foundation:** Showed deep autoencoders for dimensionality reduction
- **Citations:** 20,000+ | **Impact:** Demonstrated power of deep architectures
- **Link:** Science

25. Understanding the difficulty of training deep feedforward neural networks | Glorot & Bengio | 2010[9]

- **Theory:** Weight initialization and training dynamics
- **Citations:** 15,000+ | **Impact:** Xavier initialization still used today
- **Link:** Paper

👁 COMPUTER VISION MILESTONES *(Visual Intelligence Breakthroughs)*

26. You Only Look Once: Unified, Real-Time Object Detection (YOLO) | Redmon et al. | 2015[5]

- **Revolution:** Real-time object detection in single pass
- **Citations:** 40,000+ | **Impact:** Enabled real-time computer vision
- **Link:** arXiv | Code

27. Rich feature hierarchies for accurate object detection (R-CNN) | Girshick et al. | 2014[5]

- **Foundation:** Region-based object detection
- **Citations:** 25,000+ | **Impact:** Started modern object detection

- **Link:** arXiv | Code

28. Faster R-CNN: Towards Real-Time Object Detection | Ren et al. | 2015[5]

- **Evolution:** Unified detection pipeline with Region Proposal Networks
- **Citations:** 35,000+ | **Impact:** Standard for two-stage detection
- **Link:** arXiv | Code

29. U-Net: Convolutional Networks for Biomedical Image Segmentation | Ronneberger et al. | 2015[5]

- **Architecture:** Skip connections for dense prediction tasks
- **Citations:** 60,000+ | **Impact:** Standard for segmentation tasks
- **Link:** arXiv | Code

30. An Image is Worth 16×16 Words: Transformers for Image Recognition (ViT) | Dosovitskiy et al. | 2020[10]

- **Revolution:** Applied Transformers to computer vision
- **Citations:** 15,000+ | **Impact:** Started vision transformer era
- **Link:** arXiv | Code

31. Feature Pyramid Networks for Object Detection | Lin et al. | 2016[5]

- **Architecture:** Multi-scale feature representation
- **Citations:** 15,000+ | **Impact:** Improved detection across scales
- **Link:** arXiv | Code

32. Mask R-CNN | He et al. | 2017[5]

- **Extension:** Instance segmentation framework
- **Citations:** 20,000+ | **Impact:** State-of-the-art instance segmentation
- **Link:** arXiv | Code

NATURAL LANGUAGE PROCESSING EVOLUTION *(Language Understanding Journey)*

33. Distributed Representations of Words and Phrases (Word2Vec) | Mikolov et al. | 2013[1]

- **Foundation:** Dense word embeddings, semantic relationships
- **Citations:** 50,000+ | **Impact:** Modern NLP foundation
- **Link:** arXiv | Code

34. Sequence to Sequence Learning with Neural Networks | Sutskever et al. | 2014[1]

- **Architecture:** Encoder-decoder for sequence transduction
- **Citations:** 25,000+ | **Impact:** Foundation for machine translation

- **Link:** arXiv | Code

35. GloVe: Global Vectors for Word Representation | Pennington et al. | 2014 [1]

- **Method:** Global statistics for word embeddings
- **Citations:** 30,000+ | **Impact:** Alternative to Word2Vec
- **Link:** Paper | Code

36. ELMo: Deep contextualized word representations | Peters et al. | 2018 [1]

- **Innovation:** Context-dependent embeddings
- **Citations:** 12,000+ | **Impact:** Bridged to modern contextual models
- **Link:** arXiv | Code

37. Improving Language Understanding by Generative Pre-Training (GPT-1) | Radford et al. | 2018 [1]

- **Paradigm:** Unsupervised pre-training for language understanding
- **Citations:** 8,000+ | **Impact:** Started GPT lineage
- **Link:** Paper | Code

38. Language Models are Unsupervised Multitask Learners (GPT-2) | Radford et al. | 2019 [1]

- **Scaling:** Larger model, emergent capabilities
- **Citations:** 12,000+ | **Impact:** Showed scaling potential
- **Link:** Paper | Code

39. RoBERTa: A Robustly Optimized BERT Pretraining Approach | Liu et al. | 2019 [1]

- **Optimization:** Improved BERT training recipe
- **Citations:** 8,000+ | **Impact:** Better BERT-style models
- **Link:** arXiv | Code

GENERATIVE MODELS *(Synthesis & Creation)*

40. Unsupervised Representation Learning with Deep Convolutional GANs (DCGAN) | Radford et al. | 2015 [11]

- **Architecture:** Stabilized GAN training with convolutions
- **Citations:** 15,000+ | **Impact:** Made GANs practical
- **Link:** arXiv | Code

41. Conditional Generative Adversarial Nets | Mirza & Osindero | 2014 [11]

- **Extension:** Controlled generation with conditioning
- **Citations:** 12,000+ | **Impact:** Enabled controllable synthesis
- **Link:** arXiv | Code

42. Wasserstein GAN | Arjovsky et al. | 2017 [11]

- **Theory:** Improved training stability with Earth Mover's Distance
- **Citations:** 8,000+ | **Impact:** More stable GAN training
- **Link:** arXiv | Code

43. A Style-Based Generator Architecture (StyleGAN) | Karras et al. | 2018 [11]

- **Innovation:** Disentangled style control in generation
- **Citations:** 6,000+ | **Impact:** High-quality controllable faces
- **Link:** arXiv | Code

44. Denoising Diffusion Probabilistic Models | Ho et al. | 2020 [7]

- **New Paradigm:** Diffusion process for generation
- **Citations:** 8,000+ | **Impact:** Alternative to GANs, enabled DALL-E 2
- **Link:** arXiv | Code

45. Progressive Growing of GANs | Karras et al. | 2017 [11]

- **Training:** Progressive resolution for stable training
- **Citations:** 4,000+ | **Impact:** High-resolution image generation
- **Link:** arXiv | Code

46. CycleGAN: Unpaired Image-to-Image Translation | Zhu et al. | 2017 [11]

- **Innovation:** Translation without paired data
- **Citations:** 15,000+ | **Impact:** Artistic style transfer, domain adaptation
- **Link:** arXiv | Code

□ REINFORCEMENT LEARNING *(Learning Through Interaction)*

47. Asynchronous Methods for Deep Reinforcement Learning (A3C) | Mnih et al. | 2016 [4]

- **Architecture:** Asynchronous actor-critic methods
- **Citations:** 8,000+ | **Impact:** Scalable policy gradient methods
- **Link:** arXiv | Code

48. Proximal Policy Optimization (PPO) | Schulman et al. | 2017 [4]

- **Algorithm:** Simple, effective policy gradient method
- **Citations:** 5,000+ | **Impact:** Most popular RL algorithm today
- **Link:** arXiv | Code

49. Trust Region Policy Optimization (TRPO) | Schulman et al. | 2015 [4]

- **Theory:** Theoretical guarantees for policy updates

- **Citations:** 4,000+ | **Impact:** Principled policy improvement
- **Link:** arXiv | Code

50. Deep Deterministic Policy Gradient (DDPG) | Lillicrap et al. | 2015 [4]

- **Method:** Actor-critic for continuous control
- **Citations:** 8,000+ | **Impact:** Continuous action spaces
- **Link:** arXiv | Code

51. Mastering the Game of Go with Deep Neural Networks (AlphaGo) | Silver et al. | 2016 [1]

- **Breakthrough:** Superhuman performance in Go
- **Citations:** 12,000+ | **Impact:** Proved RL potential in complex games
- **Link:** Nature | Code

52. Rainbow: Combining Improvements in Deep Reinforcement Learning | Hessel et al. | 2017 [4]

- **Integration:** Combined multiple DQN improvements
- **Citations:** 3,000+ | **Impact:** State-of-the-art value-based RL
- **Link:** arXiv | Code

⚙ OPTIMIZATION & TRAINING *(Making Learning Work)*

53. Batch Normalization: Accelerating Deep Network Training | Ioffe & Szegedy | 2015 [9]

- **Technique:** Normalizing layer inputs for stable training
- **Citations:** 40,000+ | **Impact:** Standard in most architectures
- **Link:** arXiv | Code

54. Dropout: A Simple Way to Prevent Neural Networks from Overfitting | Srivastava et al. | 2014 [1]

- **Regularization:** Random neuron dropping during training
- **Citations:** 50,000+ | **Impact:** Standard regularization technique
- **Link:** Paper | Code

55. Layer Normalization | Ba et al. | 2016 [9]

- **Alternative:** Layer-wise normalization for RNNs/Transformers
- **Citations:** 8,000+ | **Impact:** Essential for Transformer training
- **Link:** arXiv | Code

56. On the Convergence of Adam and Beyond | Reddi et al. | 2019 [12]

- **Theory:** Fixed Adam convergence issues
- **Citations:** 2,000+ | **Impact:** AMSGrad and improved optimizers

- **Link:** arXiv | Code

🚀 MODERN LLMs & FOUNDATION MODELS *(Current AI Revolution)*

57. LLaMA: Open and Efficient Foundation Language Models | Touvron et al. | 2023[7]

- **Open Source:** Competitive open-source LLM
- **Citations:** 8,500+ | **Impact:** Democratized access to powerful LLMs
- **Link:** arXiv | Code

58. GPT-4 Technical Report | OpenAI | 2023[7]

- **Milestone:** Multimodal large language model
- **Citations:** 3,400+ | **Impact:** Current state-of-the-art general AI
- **Link:** arXiv | API

59. PaLM: Scaling Language Modeling with Pathways | Chowdhery et al. | 2022[7]

- **Scale:** 540B parameter model with emergent abilities
- **Citations:** 2,000+ | **Impact:** Demonstrated benefits of extreme scale
- **Link:** arXiv

60. InstructGPT: Training language models to follow instructions | Ouyang et al. | 2022[7]

- **Alignment:** Human feedback for better instruction following
- **Citations:** 4,000+ | **Impact:** Foundation for ChatGPT
- **Link:** arXiv

61. Constitutional AI: Harmlessness from AI Feedback | Bai et al. | 2022[7]

- **Safety:** AI system training itself to be less harmful
- **Citations:** 1,500+ | **Impact:** Scalable AI alignment approach
- **Link:** arXiv

62. Flamingo: a Visual Language Model for Few-Shot Learning | Alayrac et al. | 2022[7]

- **Multimodal:** Few-shot learning with vision and language
- **Citations:** 2,000+ | **Impact:** Advanced multimodal reasoning
- **Link:** arXiv

🌐 MULTIMODAL & CROSS-MODAL *(Beyond Single Modalities)*

63. CLIP: Learning Transferable Visual Representations from Natural Language | Radford et al. | 2021[10]

- **Innovation:** Joint vision-language representation learning
- **Citations:** 8,000+ | **Impact:** Foundation for multimodal AI

- **Link:** arXiv | Code

64. DALL-E: Creating Images from Text | Ramesh et al. | 2021 [1]

- **Breakthrough:** High-quality text-to-image generation
- **Citations:** 3,000+ | **Impact:** Demonstrated multimodal generation
- **Link:** arXiv | API

65. DALL-E 2: Hierarchical Text-Conditional Image Generation | Ramesh et al. | 2022 [7]

- **Evolution:** Improved text-to-image with diffusion
- **Citations:** 2,000+ | **Impact:** Mainstream AI art generation
- **Link:** arXiv | API

66. Segment Anything | Kirillov et al. | 2023 [7]

- **Foundation:** Universal image segmentation model
- **Citations:** 5,300+ | **Impact:** Zero-shot segmentation capabilities
- **Link:** arXiv | Code

67. LLaVA: Visual Instruction Tuning | Liu et al. | 2023 [7]

- **Innovation:** Large multimodal model for visual chat
- **Citations:** 2,800+ | **Impact:** Multimodal conversational AI
- **Link:** arXiv | Code

THEORETICAL BREAKTHROUGHS *(Understanding Why Things Work)*

68. The Lottery Ticket Hypothesis | Frankle & Carbin | 2018 [1]

- **Theory:** Sparse subnetworks can match full network performance
- **Citations:** 3,000+ | **Impact:** Network pruning and efficiency
- **Link:** arXiv | Code

69. Deep Double Descent | Nakkiran et al. | 2019 [1]

- **Phenomenon:** Performance improves again after classical overfitting
- **Citations:** 1,500+ | **Impact:** Rethinking bias-variance tradeoff
- **Link:** arXiv

70. Neural Tangent Kernel | Jacot et al. | 2018 [1]

- **Theory:** Infinite-width neural networks as kernel methods
- **Citations:** 2,500+ | **Impact:** Theoretical understanding of training
- **Link:** arXiv | Code

71. Understanding deep learning requires rethinking generalization | Zhang et al. | 2016 [1]

- **Theory:** Deep networks can memorize random labels
- **Citations:** 8,000+ | **Impact:** Challenged traditional ML theory
- **Link:** arXiv

72. Grokking: Generalization Beyond Overfitting | Power et al. | 2022 [10]

- **Phenomenon:** Sudden generalization long after overfitting
- **Citations:** 500+ | **Impact:** New understanding of learning dynamics
- **Link:** arXiv

🚀 RECENT BREAKTHROUGHS (2022-2025) *(Cutting-Edge Developments)*

73. Mamba: Linear-Time Sequence Modeling with Selective State Spaces | Gu & Dao | 2023 [13]

- **Architecture:** Alternative to Transformers with linear complexity
- **Citations:** 2,000+ | **Impact:** Potential Transformer successor
- **Link:** arXiv | Code

74. Direct Preference Optimization | Rafailov et al. | 2023 [14]

- **Training:** Simplified RLHF without separate reward model
- **Citations:** 1,000+ | **Impact:** More efficient alignment training
- **Link:** arXiv | Code

75. Tree of Thoughts: Deliberate Problem Solving | Yao et al. | 2023 [7]

- **Reasoning:** Structured exploration for complex problems
- **Citations:** 800+ | **Impact:** Enhanced LLM reasoning capabilities
- **Link:** arXiv | Code

76. Retrieval-Augmented Generation for Knowledge-Intensive NLP | Lewis et al. | 2020 [7]

- **Architecture:** Combining retrieval with generation
- **Citations:** 4,000+ | **Impact:** Foundation for modern RAG systems
- **Link:** arXiv | Code

77. Chain-of-Thought Prompting Elicits Reasoning | Wei et al. | 2022 [7]

- **Method:** Step-by-step reasoning in language models
- **Citations:** 3,000+ | **Impact:** Improved LLM problem-solving
- **Link:** arXiv

□ ESSENTIAL READING STRATEGY

Priority Levels:

- □ **MUST-KNOW (Papers 1-15):** Essential for any AI practitioner
- □ **FOUNDATIONS (Papers 16-25):** Historical context and theory
- □ **DOMAIN SPECIFIC (Papers 26-72):** Pick your specialization
- □ **CUTTING EDGE (Papers 73-77):** Stay current with latest developments

Reading Roadmap:

1. **Week 1-2:** Read all 15 MUST-KNOW papers
2. **Week 3-4:** Foundations (16-25) for historical context
3. **Month 2:** Deep dive into your domain of interest
4. **Month 3:** Explore other domains for breadth
5. **Ongoing:** Stay updated with recent breakthroughs

Implementation Strategy:

- **For each paper:** Read, understand, implement key concepts
- **Build portfolio:** 3-5 implementations from different domains
- **Join communities:** Follow authors on Twitter, join Discord servers
- **Stay current:** Follow ArXiv, attend conferences (NeurIPS, ICML, ICLR)

Success Metrics:

- **Beginner → Intermediate:** Understand and can explain all 15 must-knows
- **Intermediate → Advanced:** Deep knowledge in 1-2 domains, broad knowledge across all
- **Advanced → Expert:** Contributing to field, citing and building on these foundations

This list represents the **canonical knowledge** every serious AI practitioner should possess. These papers have shaped the field and continue to influence current research and development.

⁂

1. https://www.reddit.com/r/MachineLearning/comments/zetvmd/d_if_you_had_to_pick_1020_significant_papers_that/
2. https://www.nature.com/articles/nature14236
3. https://arxiv.org/abs/1412.6980
4. https://github.com/RPC2/DRL_paper_summary
5. https://viso.ai/deep-learning/computer-vision-papers/
6. https://blog.researchpal.co/academic-research/the-10-most-influential-ai-research-papers-of-all-time/

7. https://www.zeta-alpha.com/post/analyzing-the-homerun-year-for-llms-the-top-100-most-cited-ai-papers-in-2023-with-all-medals-for-o
8. https://papers.baulab.info
9. https://www.doradolist.com/papers/21-most-cited-machine-learning-papers
10. https://www.reddit.com/r/MachineLearning/comments/16ij18f/d_the_ml_papers_that_rocked_our_world_20202023/
11. https://www.reddit.com/r/learnmachinelearning/comments/mm8a4j/mustread_papers_on_generative_adversarial/
12. https://ceur-ws.org/Vol-3742/paper17.pdf
13. https://arxiv.org/abs/2312.00752
14. https://github.com/SarahRastegar/Best-Papers-Top-Venues
15. https://www.linkedin.com/pulse/top-10-breakthrough-computer-vision-technologies-shaping-2024-v1igf
16. https://en.wikipedia.org/wiki/Deep_learning
17. https://machinelearningmastery.com/5-breakthrough-machine-learning-research-papers-already-in-2025/
18. https://github.com/floodsung/Deep-Learning-Papers-Reading-Roadmap
19. https://machinelearningmastery.com/5-most-influential-machine-learning-papers-2024/
20. https://www.dataversity.net/brief-history-deep-learning/
21. https://cvpr.thecvf.com/Conferences/2025/News/Awards_Press
22. https://github.com/terryum/awesome-deep-learning-papers
23. https://www.sciencedirect.com/science/article/pii/S1568494625006891
24. https://github.com/youssefHosni/Weekly-Top-Computer-Vision-Papers
25. https://www.techrxiv.org/users/778163/articles/912792-deep-learning-historical-overview-from-inception-to-actualization-models-applications-and-future-trends
26. https://huggingface.co/papers/trending
27. https://www.reddit.com/r/MachineLearning/comments/1i39iuh/d_recommendations_of_noteworthy_ai_papers_for/
28. https://digitalcommons.wku.edu/cgi/viewcontent.cgi?article=4558&context=theses
29. http://dtransposed.github.io/blog/2018/12/31/Best-of-GANs-2018-(Part-1-out-of-2)/
30. https://www.mit.edu/~jgabbard/assets/18085_Project_final.pdf
31. https://neptune.ai/blog/6-gan-architectures
32. https://docs.pytorch.org/tutorials/intermediate/reinforcement_q_learning.html
33. https://arxiv.org/pdf/1412.6980.pdf
34. https://pmc.ncbi.nlm.nih.gov/articles/PMC8921387/
35. https://arxiv.org/abs/1901.00137
36. https://proceedings.neurips.cc/paper/2020/file/f3f27a324736617f20abbf2ffd806f6d-Paper.pdf
37. https://www.sciencedirect.com/science/article/pii/S2667096820300045
38. https://github.com/nightrome/really-awesome-gan

39. https://www.reddit.com/r/reinforcementlearning/comments/1is773d/must_read_papers_for_reinforcement_learning/
40. https://www.ewadirect.com/proceedings/ace/article/view/2366
41. https://arxiv.org/vc/arxiv/papers/2005/2005.00065v1.pdf
42. https://spinningup.openai.com/en/latest/spinningup/keypapers.html