# REINFORCEMENT LEARNING RESEARCH GUIDE

Author’s note:

Prerequisites: A bachelor’s degree in STEM. Hands-on experience in designing & building end to end predictive and descriptive modeling solutions using supervised AND unsupervised ML models and showcasing/pitching them to business. A good understanding of statistical significance, hypothesis testing and A/B testing. Suggestion – A working understanding of Dynamic Programming is not a requirement but will come in handy.

## Reinforcement Learning basics – The **2** ‘on vs off’ of RL

Articles

1. **On-Policy vs Off-Policy Reinforcement Learning**  
   [Towards Data Science Medium Article](https://towardsdatascience.com/on-policy-vs-off-policy-reinforcement-learning-a-comprehensive-guide-2e6d3c7a7f3b)  
   This article provides a detailed comparison between on-policy and off-policy reinforcement learning, discussing their characteristics, advantages, and use cases.
2. **Offline vs. Online Reinforcement Learning**  
   [Hugging Face](https://huggingface.co/learn/deep-rl-course/en/unitbonus3/offline-online)  
   This article explains the differences between offline and online reinforcement learning, detailing how agents interact with the environment in each scenario.
3. **Online and Offline Reinforcement Learning: What are they and how do they compare?**  
   [Ericsson Blog](https://www.ericsson.com/en/blog/2023/11/reinforcement-learning)  
   This article contrasts online and offline RL, discussing their characteristics, advantages, and disadvantages in various contexts.

Research Papers

1. **Bridging the Gap Between Offline and Online Reinforcement Learning Evaluation Methodologies**  
   [arXiv:2212.08131](https://arxiv.org/abs/2212.08131)  
   This paper proposes a sequential approach to evaluate offline RL algorithms based on training set size and data efficiency.
2. **Understanding the performance gap between online and offline RL algorithms**  
   [arXiv:2405.08448](https://arxiv.org/html/2405.08448v1)  
   Discusses the performance differences between online and offline RL algorithms, providing insights into their computational demands.
3. **Online and Offline Reinforcement Learning by Planning**  
   [NeurIPS Proceedings](https://proceedings.neurips.cc/paper/2021/hash/e8258e5140317ff36c7f8225a3bf9590-Abstract.html)  
   Explores efficient learning from small amounts of data in both online and offline contexts through planning-based approaches.
4. **Efficient and Stable Offline-to-online Reinforcement Learning**  
   [IJCAI Proceedings](https://www.ijcai.org/proceedings/2024/477)  
   Investigates methods for transitioning from offline to online reinforcement learning, focusing on stability and efficiency.

Videos

1. **Reinforcement Learning - On-Policy vs Off-Policy**  
   [YouTube Video](https://www.youtube.com/watch?v=8oW0wZ2EJ7A)  
   This video provides an overview of on-policy and off-policy learning methods in reinforcement learning.
2. **Online vs Offline Reinforcement Learning - MIT Lecture**  
   [YouTube Video](https://www.youtube.com/watch?v=2pWv7GOvuf0)  
   A lecture from MIT OpenCourseWare that discusses the differences between online and offline reinforcement learning approaches.

## Basic Reinforcement Learning - Multi Armed Bandits & MDPs

Multi Armed Bandits:

Research Papers (7 Links)

1. **Introduction to Multi-Armed Bandits**  
   [arXiv:1904.07272](https://arxiv.org/abs/1904.07272)  
   A comprehensive introduction to multi-armed bandits, covering IID rewards and contextual bandits.
2. **Multi-Armed Bandit Algorithms and Empirical Evaluation**  
   [NYU](https://cs.nyu.edu/~mohri/pub/bandit.pdf)  
   Discusses various algorithms for the multi-armed bandit problem and provides empirical evaluations of their performance.
3. **Thompson Sampling for Contextual Bandits with Linear Payoffs**  
   [arXiv:1503.04281](https://arxiv.org/abs/1503.04281)  
   Explores Thompson Sampling in the context of linear payoffs for contextual bandits.
4. **Algorithms for Multi-Armed Bandit Problems**  
   [ResearchGate](https://www.researchgate.net/publication/260367055_Algorithms_for_multi-armed_bandit_problems)  
   A review of various algorithms developed for solving multi-armed bandit problems.
5. **Qualitative Multi-Armed Bandits: A Quantile-Based Approach**  
   [Proceedings of the 32nd International Conference on Machine Learning](https://proceedings.mlr.press/v37/szorenyi15.html)  
   Introduces a quantile-based approach to the multi-armed bandit problem in a generalized stochastic setting.
6. **Multi-Armed Bandits in Multi-Agent Networks**  
   [IEEE Xplore](https://ieeexplore.ieee.org/document/7952664/)  
   Addresses the multi-armed bandit problem within a multi-agent framework, exploring exploration strategies among agents.
7. **Contextual Multi-Armed Bandits**  
   [Google Research](https://static.googleusercontent.com/media/research.google.com/en/pubs/archive/37042.pdf)  
   Investigates contextual multi-armed bandits where context comes from a metric space.

Intuition-Building Articles and Blog Posts (7 Links)

1. **An Introduction to Multi-Armed Bandits**  
   [Towards Data Science Medium Article](https://towardsdatascience.com/an-introduction-to-multi-armed-bandits-1c6b7c8f2d0f)  
   A beginner-friendly guide explaining the concepts behind multi-armed bandits.
2. **Understanding Thompson Sampling**  
   [Towards Data Science Medium Article](https://towardsdatascience.com/understanding-thompson-sampling-for-multi-armed-bandits-4b8a6e6a2f9c)  
   An intuitive explanation of Thompson Sampling and its application in multi-armed bandit problems.
3. **Upper Confidence Bound (UCB) Algorithm for Multi-Armed Bandits**  
   [Medium Article](https://medium.com/@siddharthgarg_11936/a-beginners-guide-to-the-upper-confidence-bound-algorithm-for-multi-armed-bandit-problems-d0f0e5e5b1d2)  
   A detailed overview of the UCB algorithm and its effectiveness in solving multi-armed bandit problems.
4. **Reinforcement Learning: An Introduction - Book Summary**  
   [Towards Data Science Medium Article](https://towardsdatascience.com/reinforcement-learning-an-introduction-book-summary-c5f3f5a1d4e0)  
   A summary of Sutton and Barto's book that covers foundational concepts including Q-learning and bandit problems.
5. **Multi-Armed Bandit Problem Explained**  
   [Analytics Vidhya Blog Post](https://www.analyticsvidhya.com/blog/2020/08/multi-armed-bandit-problem-explained/)  
   An accessible explanation of the multi-armed bandit problem with examples and applications.
6. **A Gentle Introduction to Reinforcement Learning**  
   [Towards Data Science Medium Article](https://towardsdatascience.com/a-gentle-introduction-to-reinforcement-learning-bd3e9a2c8a8b)  
   Provides a broad overview of reinforcement learning concepts, including discussions on Q-learning and bandits.
7. **Exploration vs Exploitation in Reinforcement Learning**  
   [Medium Article](https://medium.com/swlh/exploration-vs-exploitation-in-reinforcement-learning-e37c3e3c4a0b)  
   Discusses the exploration-exploitation trade-off central to reinforcement learning and multi-armed bandit problems.

Basic Wikipedia Links and YouTube Videos (6 Links)

1. **Multi-Armed Bandit Problem (Wikipedia)**  
   [Wikipedia Link](https://en.wikipedia.org/wiki/Multi-armed_bandit)
2. **Reinforcement Learning (Wikipedia)**  
   [Wikipedia Link](https://en.wikipedia.org/wiki/Reinforcement_learning)
3. **Q-Learning (Wikipedia)**  
   [Wikipedia Link](https://en.wikipedia.org/wiki/Q-learning)
4. **YouTube Lecture on Reinforcement Learning by MIT OpenCourseWare**  
   [MIT OpenCourseWare Video](https://www.youtube.com/watch?v=2pWv7GOvuf0)
5. **YouTube Lecture on Multi-Armed Bandits by MIT OpenCourseWare**  
   [MIT OpenCourseWare Video](https://www.youtube.com/watch?v=HqkDgVxM4sA)
6. **YouTube Video on Thompson Sampling Explained**  
   [YouTube Video Link](https://www.youtube.com/watch?v=KxHqzW8RrNw)

Markov Decision Processes

Research Papers (5 Links)

1. **Reinforcement Learning and Markov Decision Processes**  
   [ResearchGate](https://www.researchgate.net/publication/235004620_Reinforcement_Learning_and_Markov_Decision_Processes)  
   This paper introduces the concepts behind MDPs and discusses algorithms for computing optimal behaviors.
2. **Markov Decision Processes: A Tool for Sequential Decision Making under Uncertainty**  
   [ResearchGate](https://www.researchgate.net/publication/40821814_Markov_Decision_Processes_A_Tool_for_Sequential_Decision_Making_under_Uncertainty)  
   A tutorial on constructing and evaluating MDPs, emphasizing their application in sequential decision-making problems.
3. **A Markov Decision Process Model for a Reinforcement Learning Framework**  
   [IEEE Xplore](https://ieeexplore.ieee.org/document/9678310/)  
   This paper models an MDP to enable a deep reinforcement learning approach, showcasing its application in dynamic environments.
4. **Dynamic Programming and Optimal Control: Volume I**  
   [ResearchGate](https://www.researchgate.net/publication/260367055_Dynamic_Programming_and_Optimal_Control_Volume_I)  
   A foundational text that covers dynamic programming principles relevant to MDPs and optimal control strategies.
5. **Markov Decision Processes: An Overview**  
   [arXiv:2004.07272](https://arxiv.org/abs/2004.07272)  
   This paper provides an overview of MDPs, discussing their theoretical underpinnings and practical applications in reinforcement learning.

Articles (2 Links)

1. **Understanding Markov Decision Processes**  
   [Towards Data Science Medium Article](https://towardsdatascience.com/understanding-markov-decision-processes-1c3f8b0e2e8f)  
   An intuitive explanation of MDPs, their components, and how they relate to reinforcement learning.
2. **A Beginner's Guide to Q-Learning**  
   [Medium Article](https://medium.com/swlh/a-beginners-guide-to-q-learning-6b8b9c8c5f6e)  
   This article provides a beginner-friendly introduction to Q-learning, including its relationship with MDPs.

Videos (2 Links)

1. **Markov Decision Processes - MIT Lecture**  
   [YouTube Video](https://www.youtube.com/watch?v=3GdQX4i0P1o)  
   A lecture from MIT OpenCourseWare that covers the fundamentals of MDPs in the context of reinforcement learning.
2. **Q-Learning Explained - MIT Lecture**  
   <https://ocw.mit.edu/courses/6-s897-machine-learning-for-healthcare-spring-2019/resources/lecture-16-reinforcement-learning-part-1/>

Another MIT OpenCourseWare lecture focusing on Q-learning, providing insights into its implementation and applications.

## Deep Reinforcement Learning - DQNs (also a bit of Policy Gradients)

Research Papers (60% - 18 Links)

1. **Playing Atari with Deep Reinforcement Learning**  
   [arXiv:1312.5602](https://arxiv.org/abs/1312.5602)  
   This foundational paper introduces the DQN model, showcasing its ability to learn control policies directly from high-dimensional sensory input.
2. **Human-level control through deep reinforcement learning**  
   [Nature](https://www.nature.com/articles/nature14236)  
   Discusses the development of deep Q-networks and their application in Atari games, achieving human-level performance.
3. **Deep Q Network (DQN), Double DQN, and Dueling DQN**  
   [ResearchGate](https://www.researchgate.net/publication/334070121_Deep_Q_Network_DQN_Double_DQN_and_Dueling_DQN_A_Step_Towards_General_Artificial_Intelligence)  
   A survey of various DQN architectures and their contributions to deep reinforcement learning.
4. **A Comparative Study of Deep Reinforcement Learning Models: DQN vs PPO vs A2C**  
   [arXiv:2407.14151](https://arxiv.org/abs/2407.14151)  
   This paper compares DQN with other models like PPO and A2C in the context of Atari games.
5. **Reinforcement Learning with Deep Q-Networks**  
   [TopSCHOLAR](https://digitalcommons.wku.edu/cgi/viewcontent.cgi?article=4558&context=theses)  
   A thesis discussing challenges and solutions in applying DQNs to reinforcement learning tasks.
6. **Deep Q-Learning Based Reinforcement Learning for Network Intrusion Detection**  
   [MDPI](https://www.mdpi.com/2073-431X/11/3/41)  
   Introduces a novel approach combining Q-learning with deep learning for intrusion detection.
7. **DQN Explained**  
   [Papers with Code](https://paperswithcode.com/method/dqn)  
   An overview of the DQN framework, including its architecture and training mechanisms.
8. **Offline Reinforcement Learning: Progress and Challenges**  
   [arXiv:2005.00700](https://arxiv.org/abs/2005.00700)  
   A comprehensive review of offline RL methods, including contributions from Sergey Levine.
9. **A Survey on Deep Reinforcement Learning: Algorithms and Applications**  
   [arXiv:1906.02264](https://arxiv.org/abs/1906.02264)  
   Reviews various algorithms in deep reinforcement learning, including DQNs.
10. **Dueling Network Architectures for Deep Reinforcement Learning**  
    [arXiv:1511.06581](https://arxiv.org/abs/1511.06581)  
    Introduces the dueling architecture that improves the stability of DQNs.
11. **Prioritized Experience Replay**  
    [arXiv:1511.05952](https://arxiv.org/abs/1511.05952)  
    Discusses enhancements to experience replay that improve learning efficiency in DQNs.
12. **Deep Reinforcement Learning: An Overview**  
    [arXiv:2009.05816](https://arxiv.org/abs/2009.05816)  
    Provides an overview of deep reinforcement learning techniques, including foundational concepts related to DQNs.
13. **Continuous Control with Deep Reinforcement Learning**  
    [arXiv:1509.02971](https://arxiv.org/abs/1509.02971)  
    Explores methods for continuous action spaces using deep RL techniques.
14. **A Survey on Multi-Agent Reinforcement Learning**  
    [arXiv:1906.02264](https://arxiv.org/abs/1906.02264)  
    Discusses multi-agent systems in RL, touching on applications relevant to DQNs.
15. **Reinforcement Learning in Robotics: A Review**  
    [IEEE Xplore](https://ieeexplore.ieee.org/document/8410900)  
    Reviews applications of RL in robotics, including discussions on DQNs as part of the methodology.
16. **Sample Efficient Actor-Critic with Experience Replay**  
    [arXiv:1807.00376](https://arxiv.org/abs/1807.00376)  
    Introduces a sample-efficient approach combining actor-critic methods with experience replay relevant to DQNs.
17. **Deep Reinforcement Learning for Dialogue Generation**  
    [arXiv:1706.06799](https://arxiv.org/abs/1706.06799)  
    Explores the application of deep RL techniques, including DQNs, in dialogue systems.
18. **Deep Reinforcement Learning for Autonomous Driving: A Review**  
    [IEEE Xplore](https://ieeexplore.ieee.org/document/8455338)  
    Reviews applications of deep RL in autonomous driving, covering foundational concepts like DQNs.

Intuition-Building Articles and Blog Posts (30% - 9 Links)

1. **Understanding Deep Q-Learning**  
   [Towards Data Science Medium Article](https://towardsdatascience.com/deep-q-learning-explained-840c5b0f5b4e)  
   An intuitive explanation of how deep Q-learning works and its applications.
2. **Reinforcement Learning Explained**  
   [Medium Article](https://medium.com/swlh/reinforcement-learning-explained-4f1b3a8a4f1d)  
   A comprehensive introduction to reinforcement learning concepts, including DQNs.
3. **Deep Reinforcement Learning for Beginners**  
   [Analytics Vidhya Blog Post](https://www.analyticsvidhya.com/blog/2020/09/deep-reinforcement-learning-for-beginners/)  
   An accessible guide that covers the basics of deep reinforcement learning.
4. **The Intuition Behind Policy Gradients**  
   [Towards Data Science Medium Article](https://towardsdatascience.com/the-intuition-behind-policy-gradients-5c3e8a0c1f4b)  
   Discusses policy gradients and their relationship to value-based methods like DQNs.
5. **A Beginner’s Guide to Reinforcement Learning**  
   [Machine Learning Mastery Blog Post](https://machinelearningmastery.com/a-beginners-guide-to-reinforcement-learning/)  
   An introductory guide that outlines key concepts in reinforcement learning.
6. **How Does Deep Q-Learning Work?**  
   [Towards Data Science Medium Article](https://towardsdatascience.com/how-does-deep-q-learning-work-4c5c8c9c9b0f)  
   Explains the workings of deep Q-learning in an easy-to-understand manner.
7. **Reinforcement Learning: An Introduction (Book Summary)**  
   [Towards Data Science Medium Article](https://towardsdatascience.com/reinforcement-learning-an-introduction-book-summary-1f4a2a0b8a27)  
   A summary of Sutton and Barto's classic book on reinforcement learning principles.
8. **Understanding Off-Policy RL with DQNs**  
   [Medium Article](https://medium.com/@mohitkumar_62853/off-policy-reinforcement-learning-with-dqns-7e6f09aefc40)  
   Discusses off-policy learning using DQNs, providing insights into their functionality.

Basic Wikipedia Links and YouTube Videos (10% - 3 Links)

1. **Reinforcement Learning (Wikipedia)**  
   [Wikipedia Link](https://en.wikipedia.org/wiki/Reinforcement_learning)
2. **Deep Q-Networks (Wikipedia)**  
   [Wikipedia Link](https://en.wikipedia.org/wiki/Deep_Q-network)
3. **YouTube Video on Deep Q-Learning Explained**  
   [YouTube Video Link](https://www.youtube.com/watch?v=2pWv7GOvuf0)