**TÍTULO**

Modelo Detector de Sinais de Compra e Venda Utilizando Aprendizado por Reforço

Ex: *Deep Reinforcement Learning with Double Q-Learning - 2016*

1. **INTRODUÇÃO**
   1. Contextualização
      1. Aprendizado de máquina
         * + *Foundations of Machine Learning - 2012, The MIT Press*
           + *What Is Machine Learning - elnaqa2015*
           + Proposes a ‘learning machine’ - Alan Turing

*Computing Machinery and Intelligence - turing1950*

*Computation Beyond the Turing Limit - siegelmann1995*

* + - * + Probable World's First Machine Learning Program

*Arthur Samuel\_ Pioneer in Machine Learning - mccarthy1991*

* + - * Types of machine learning systems
        + Supervised

*A Comparative Study of Training Algorithms for Supervised Machine Learning – 2012*

*A review of supervised machine learning algorithms and their applications to ecological data - crisci2012*

* + - * + Unsupervised

Unsupervised Learning - 1999

*Building High-level Features Using Large Scale Unsupervised Learning - 2012*

Unsupervised Machine Learning for Networking\_ Techniques, Applications and Research Challenges – 2019

* + - * + Supervised/ Unsupervised

Comparing supervised and unsupervised category learning - Love2002.pdf

* + - * + Semi-Supervised

Semi-Supervised and Unsupervised Extreme Learning Machines - gaohuang2014.pdf

Chapter 22 - Semi-Supervised Learning - zhou2014.pdf

MixMatch\_ A Holistic Approach to Semi-Supervised Learning - 2019

* + - * + Reinforcement Learning

***Use the references from the reinforcement learning chapter***

* + - * + Online and Batch Learning

*On-Line Algorithms in Machine Learning - blum1998.pdf*

*Comparison of offline and real-time human activity recognition results using machine learning techniques - suto2018.pdf*

*Multi Kernel Learning with Online-Batch Optimization - orabona2012a.pdf*

* + - * + Instance-based Learning

*Instance-based learning algorithms - aha1991.pdf*

*Combining instance-based learning and logistic regression for multilabel classification – 2009*

*Instance-Based Utile Distinctions for Reinforcement Learning with Hidden State - 1995*

* + - * + Model-based Learning

***Model-based and Model-free in reinforcement learning section***

* 1. Desafio
  2. Justificativa
  3. Motivação
  4. Objetivos
  5. Organização do trabalho

1. **Trabalhos relacionados**
2. **Fundamentação teórica**
   1. redes neurais artificiais
      * + *https://www.dataversity.net/brief-history-deep-learning/*
        + *A logical calculus of the ideas immanent in nervous activity - mcculloch1943*
        + *Neural networks and physical systems with emergent collective computational abilities - 1982*
        + *On the Computational Power of Neural Nets - siegelmann1995.pdf*
        + Deep Learning
          - *Ivakhnenko, A. G. (1971). Polynomial Theory of Complex Systems - Possível primeira aparição Deep Learning*
          - *Deep learning in neural networks\_ An overview - schmidhuber2015.pdf*
      1. Neurônios: biológico e artificial
         * Artificial neural networks\_ a tutorial - jain1996.pdf
         * The Perceptron
           + *The perceptron - A probabilistic model for information storage and organization in the brain – 1958*

*Perceptrons\_ An introduction to computational geometry (1969, MIT Press) – 1969 ---- Não resolve XOR*

*In 1969, Minsky and Papert’s book Perceptrons was published. It was a harsh critique on Rosenblatt’s perceptrons. Minsky and Papert proved that perceptrons could only be trained to solve linear separable problems. For instance, one of the most dooming examples of non-linear separable problems is the exclusive OR (XOR).*

*Artificial intelligence- a general survey – Lighthill1973*

*The Lighthill report (published in 1973) was an evaluation of the current state of AI at that time written for the British Science Research Council. The report came to the conclusion that the promises of AI researchers were exaggerated: “in no part of the field have discoveries made so far produced the major impact that was then promised.”*

* + 1. Redes feedforward
       - *Learning algorithms and probability distributions in feed-forward and feed-back networks – 1987*
       - *Training feedforward networks with the Marquardt algorithm – 1994*
  + Multilayer Perceptron Feedforward
    - *On the Capabilities of Multilayer Perceptrons – 1988*
    - *A General Multilayer Perceptrons Feed Forward Neural Network Algorithm for Learning Capability - 201*
    1. Backpropagation
       - Taylor expansion of the accumulated rounding error - 1970 - Backpropagation
         * Apesar de existir era ineficiente devido o poder computacional da época, pelo menos até 1985
         * *Learning representations by back-propagating errors - 1986 Mostra que BackPropation pode ser usado para obter melhores resultados.pdf*
         * *A theoretical framework for Back-Propagation – 1988*
         * *Improving the Convergence of Back-Propagation Learning with Second-Order Methods – 1989*
         * *Theory of the backpropagation neural network – 1989*
         * *Back-propagation fails to separate – 1989*
         * *On the problem of local minima in backpropagation – 1992*
         * *Efficient BackProp*

*On derivation of MLP backpropagation from the Kelley-Bryson optimal-control gradient formula and its application*

* 1. Função de ativação
     1. Non-Linear Activation functions
     2. https://missinglink.ai/guides/neural-network-concepts/7-types-neural-network-activation-functions-right/
     3. Binary Step Function
     4. Linear Activation Function
        + *Learning Activation Functions to Improve Deep Neural Networks*
          - Non-Linear Activation Functions

*Learning Activation Functions to Improve Deep Neural Networks*

Sigmoid / Logistic

Performance Analysis of Various Activation Functions in Generalized MLP Architectures of Neural Networks – 2011

Implementation of a Sigmoid Activation function for Neural Network - 2012

Analysis of different activation functions using back propagation neural networks – 2013

TanH / Hyperbolic Tangent

Why Tanh\_ Choosing a Sigmoidal Function - 1992

ReLU (Rectified Linear Unit)

Parameterised Sigmoid and ReLU Hidden Activation Functions for DNN Acoustic Modelling - 2015

Deep Learning using Rectified Linear Units (ReLU) – 2018

Searching for Activation Functions – 2017

Research on convolutional neural network based on improved Relu piecewise activation function - lin2018

Leaky ReLU

Comparative Study of Convolution Neural Network’s Relu and Leaky-Relu Activation Functions - 2019

Parametric ReLU

Alcoholism identification via convolutional neural network based on parametric ReLU, dropout, and batch normalization - wang2018

*Softmax*

*On The Pairing Of The Softmax Activation And Cross{Entropy Penalty Functions And The Derivation Of The Softmax Activation Function - 1997*

*Simplified Hardware Implementation of the Softmax Activation Function - kouretas2019.pdf*

*ACTIVATION FUNCTIONS IN NEURAL NETWORKS - 2020*

Swish

*Swish: a Self-Gated Activation Function*

*Flatten-T Swish\_ a thresholded ReLU-Swish-like activation function for deep learning - 2018*

* 1. Optimizadores
     + - Gradient Descent variants
         * Batch gradient descent

*The general inefficiency of batch training for gradient descent learning*

* + - * + Stochastic gradient descent

*A Stochastic Approximation Method - SGD First Paper*

*Stochastic Estimation of the Maximum of a Regression Function - SGD 2*

*An overview of gradient descent optimization*

* + - * + Mini-batch gradient descent

*Better Mini-Batch Algorithms via Accelerated Gradient Methods*

* + - * Momentum
        + *Some methods of speeding up the convergence of iteration methods - Momentum*
        + *On the importance of initialization and momentum in deep learning*
      * Grandiente acelerado Nesterov
        + *A method for solving the convex programming problem with convergence rate - Nesterov Original Paper*
        + *Accelerated Distributed Nesterov Gradient Descent*
      * AdaGrad
        + *Adaptive Subgradient Methods for Online Learning and Stochastic Optimization -ADAGRAD*
      * Adadelta
        + *ADADELTA: An Adaptive Learning Rate Method*
      * RMSProp
        + *Neural Networks for Machine - Learning Lecture 6a Overview of mini-batch gradient descent - RMSProp*
      * Adam
        + *ADAM: A METHOD FOR STOCHASTIC OPTIMIZATION*
      * ***To summarize, RMSProp, AdaDelta and Adam are very similar algorithm and since Adam was found to slightly outperform RMSProp, Adam is generally chosen as the best overall choice.***
    1. Momento
    2. Gradiente acelerado de Nesterov
  1. Overfitting e Underfitting
     + *Machine Learning Basics - Overfitting e Underfitting – Slide*
     + *Process mining\_ a two-step approach to balance between underfitting and overfitting - 2010*
     + *Methods To Avoid Over-fitting And Under-fitting In Supervised Machine Learning (comparative Study) – 2015*
     + *A pruning based method to learn both weights and connections for LSTM – 2015*
     + *Overfitting and Underfitting Analysis for Deep Learning Based End-to-end Communication Systems - 2019*
     + *Study of Dependency on number of LSTM units for Character based Text Generation models - chakraborty2020.pdf*
     + Early-Stopping
       - Early Stopping - But When - prechelt1998.pdf
       - Training Recurrent Answering Units with Joint Loss Minimization for VQA - 2016
       - HINDSIGHT\_ An R-Based Framework Towards Long Short Term Memory (LSTM) Optimization – 2018Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems - Aurélien Géron – 2019
       - https://keras.io/api/callbacks/early\_stopping/
  2. Regularization
     + - *Networks for Approximation and Learning - regularization – 1990*
       - *Feature selection, L1 vs. L2 regularization, and rotational invariance - ng2004*
       - *On Manifold Regularization - TR-2004-05*
       - *Recurrent Neural Network Regularization - 2015*
       - *Regularizing and Optimizing LSTM Language Models – 2017*
       - *What is the Effect of Importance Weighting in Deep Learning - 2019*
       - Dropout
         * *Dropout Training as Adaptive Regularization – 2013*
         * *The Implicit and Explicit Regularization Effects of Dropout – 2020*
         * *Adversarial Dropout Regularization – 2017*
       - Data Augmentation
         * The Art of Data Augmentation - vandyk2001
         * Understanding Data Augmentation for Classification\_ When to Warp - wong2016
         * Data Augmentation for Recognition of Handwritten Words and Lines Using a CNN-LSTM Network - wigington2017.pdf
         * Data Augmentation and Dense-LSTM for Human Activity Recognition using WiFi Signal - zhang2020
       - Regularization and Optimizing
         * *Regularizing and Optimizing LSTM Language Models – 2017*
         * <https://en.wikipedia.org/wiki/Vanishing_gradient_problem#Long_short-term_memory>
  3. Aprendizado por reforço
     + - [*http://incompleteideas.net/book/first/ebook/node12.html*](http://incompleteideas.net/book/first/ebook/node12.html)
* Foundations
  + *An Example of Statistical Investigation of the Text Eugene Onegin Concerning the Connection of Samples in Chains - Markov Chains 1913 - Translate 2007*
  + *A Mathematical Theory of Communication – 1948*
  + *Finding Structure in Reinforcement Learning - NIPS-1994*
  + *Learning to predict by the methods of temporal difference - Sutton1988.pdf*
  + *Learning from delayed reward - 1989.pdf*
  + *Reinforcement Learning for Robots Using Neural Networks - 1993*
  + *Reinforcement Learning and the Reward Engineering Principle - 2014*
  + *From Reinforcement Learning to Optimal Control: A unified framework for sequential decisions - 2019.pdf*
  + *Holistic Reinforcement Learning - The Role of Structure and Attention - radulescu2019*
  + *Deep Reinforcement Learning for Dexterous Manipulation with Concept Networks – 2017*
    - * **State, Action, Reward, Discount** – Concepts of Dynamic programming
  + Estrutura
    - Estado, agente, recompensa e desconto
  1. Processo de decisão de Markov
     1. A equação de Bellman
        + The Bellman equation

*On the role of dynamic programming in statistical communication theory - bellman 1957*

*Dynamic programming and stochastic control processes - bellman1958*

***introduced the discrete stochastic version of the optimal control problem known as Markovian decision processes (MDPs)\*\*\****

*On adaptive control processes - bellman1959.pdf*

*Applied dynamic programming – 1962*

* + - * *Planning*
        + *Integrated Architectures for Learning, Planning, and Reacting Based on Approximating Dynamic Programming - sutton1990.pdf*

Diagrama, Diagrama de Venn

Descrição gerada automaticamente

* Markov Decision Process
  + *Dynamic Programming - rhoward1960*
    - ***devised the policy iteration method for MDPs. All of these are essential elements underlying the theory and algorithms of modern reinforcement learning.***
  + *Means and variances of time averages in Markovian environments - 1987.pdf*
  + *Numerical Transient Analysis Of Markov Models – 1988*
  + *Transient analysis of cumulative measures of markov model behavior – 1989*
  + *Hierarchical Control and Learning for Markov Decision Processes - 1998 – 1990*
  + *A survey of algorithmic methods for partially observed Markov decision processes - lovejoy1991*
  + *A Survey of Applications of Markov Decision Processes - 1991*
  + *Markov and the Birth of Chain Dependence Theory – 1996*
  + *Markov Decision Processes\_ Concepts and Algorithms – 2009 \*\*\**
  + Markov reward process
    - *Markov and Markov reward model transient analysis - An overview of numerical approaches - reibman1989*
    - *Performability Analysis Using Semi-Markov Reward Processes – 1990*
    - *Automated Generation And Analysis Of Markov Reward Models Using Stochastic Reward Nets - 1993*
    - *Markov Reward Models and Markov Decision Processes in Discrete and Continuous Time-Performance Evaluation and Optimization – gouberman2014*
    - **Living Penalty**
    1. Backpropagation
       - *Reinforcement learning by backpropagation through an LSTM model-critic – 2007*
       - *A Gentle Tutorial of Recurrent Neural Network with Error Backpropagation – 2016*
       - *Comparison of Predictive Algorithms- Backpropagation, SVM, LSTM and Kalman Filter for Stock Market - 2019*
       - *Dense Recurrent Neural Networks for Accelerated MRI- History-Cognizant Unrolling of Optimization Algorithms - hosseini2020*
       - Backpropagation Through Time
         * Recurrent Neural Networks Trained With Backpropagation Through Time Algorithm to Estimate Nonlinear Load Harmonic Currents - mazumdar2008
       - Truncated Backpropagation Through Time
         * *Unbiasing Truncated Backpropagation Through Time – 2017*
         * *On Training Recurrent Networks with Truncated Backpropagation Through time in Speech Recognition - tang2018*
         * *Adaptively Truncating Backpropagation Through Time to Control Gradient Bias - 2020*
    2. Política
       - Iteração de política
       - Iteração de valor
       - Gradiente de Política Natural
       - Monte Carlo
         * Policy

*A policy-blending formalism for shared control - dragan2013*

*Policy Gradient Methods for Reinforcement Learning with Function Approximation – 1999*

*Composite Task-Completion Dialogue Policy Learning via Hierarchical Deep Reinforcement Learning – 2017*

Fast Policy Learning through Imitation and Reinforcement – 2018

*Improving RTS Game AI by Supervised Policy Learning, Tactical Search, and Deep Reinforcement Learning - barriga2019.pdf*

**Policy Gradients**

*Reinforcement Learning Through Gradient Descent – 1999*

*Gradient Descent for General Reinforcement Learning – 1999*

*Policy gradient methods for reinforcement learning with function approximation - 2000*

*Policy Gradient Reinforcement Learning for Fast Quadrupedal Locomotion - kohl2004*

*Restricted gradient-descent algorithm for value-function approximation in reinforcement learning – 2007*

*Restricted gradient-descent algorithm for value-function approximation in reinforcement learning - damottasallesbarreto2008*

*The Optimal Reward Baseline for Gradient·Based Reinforcement Learning – 201*

* + 1. Model-based and model-free learning
       - *Model-based*
         * *Model-based learning for mobile robot navigation from the dynamical systems perspective - tani1996*
         * *Introduction to model-based teaching and learning in science education - gobert2000*
         * *Multiple Model-Based Reinforcement Learning - doya2002*
         * *Model-Based Reinforcement Learning for Atari - 2020*
       - *Model-Free Learning*
         * *Model-free Q-learning designs for linear discrete-time zero-sum games with application to H-infinity control - al-tamimi2007*
         * *Learning model-free robot control by a Monte Carlo EM algorithm - vlassis2009*
         * *Generalized TD Learning - 2011*
         * *Model-Free reinforcement learning with continuous action in practice - degris2012*
         * *Neural Computations Underlying Arbitration between Model-Based and Model-free Learning - lee2014*
         * *Model-Based Value Estimation for Efficient Model-Free Reinforcement Learning – 2018*
  1. Q-Learning]
     + - *Quality of certain action in a given state*
       - *Watkins-Dayan1992\_Article\_Q-learning*
     1. Deep Q-Learning
        + *Experiência de Replay*
          - *Prioritized Experience Replay – 2015*
          - *Distributed Prioritized Experience Replay – 2018*
          - *Experience Replay Optimization – 2019*
     2. Gradiente de política
     3. Eficiência da Amostra
     4. On-policy learning vs. Off-policy learning
     5. Time difference learning
     6. Forward model
        + *Self-improving reactive agents based on reinforcement learning, planning and teaching – 1992*
        + *Ants and Reinforcement Learning\_ A Case Study in Routing in Dynamic Networks – 1998*
        + *Forward-Backward Reinforcement Learning – 2018*
     7. Backpropation model
        + *On the use of backpropagation in associative reinforcement learning - williams1988.pdf*
        + *Input Generalization in Delayed Reinforcement Learning\_ An Algorithm and Performance Comparisons – 1991*
        + *Reinforcement learning by backpropagation through an LSTM model-critic – 2007*
  2. Desafios
     + - *Introduction- The Challenge of Reinforcement Learning – 1992*
       - *Learning agents for uncertain environments (extended abstract) – 1998*
       - *Constrained Markov decision processes – 1999*
       - *Challenges of Real-World Reinforcement Learning – 2019*
       - *An empirical investigation of the challenges of real-world reinforcement learning – 2020*
     1. Problema de atribuição de crédito
        + *TEMPORAL CREDIT ASSIGNMENT IN REINFORCEMENT LEARNING - 1984.pdf*
        + *Unifying Temporal and Structural Credit Assignment Problems - 2004.pdf*
        + *Solving the credit assignment problem: explicit and implicit learning of action sequences with probabilistic outcomes - fu2007.pdf*
     2. Sparce rewards
        + *Reinforcement Learning with Unsupervised Auxiliary Tasks - 2016*
        + *Leveraging Demonstrations for Deep Reinforcement – 2017*
        + *Overcoming Exploration in Reinforcement Learning with Demonstrations - nair2018*
        + *Hindsight Experience Replay - 2018*
        + *Deep-Reinforcement-Learning-Based Autonomous UAV Navigation With Sparse Rewards - wang2020*
     3. Reward shape
        + Reward shaping for reinforcement learning by emotion expressions - hwang2014
     4. O problema de alinhamento
        + The Alignment Problem for Bayesian History-Based Reinforcement Learners - everitt2018.pdf
        + Using deep reinforcement learning approach for solving the multiple sequence alignment problem – 2019
     5. The exploration vs exploitation
        + Exploration and Exploitation in Organizational Learning - March\_1991.pdf
        + The neuroscientific foundations of the exploration−exploitation dilemma - 2010
        + Reinforcement learning\_ exploration–exploitation dilemma in multi-agent foraging task - yogeswaran2012.pdf
        + The Exploration-Exploitation Dilemma\_ A Multidisciplinary Framework - berger-tal2014.pdf
        + MULEX\_ Disentangling Exploitation from Exploration in Deep RL – 2017
        + Solution *epsilon* Gradient
     6. Curiosity driven exploration
        + *Agent can stuck in local minimum with small, recurring rewards*
        + *Curiosity-driven Exploration by Self-supervised Prediction – 2017*
        + *Adaptive ε-Greedy Exploration in Reinforcement Learning Based on Value Differences – 2010 \*\*\**
        + *Classes of Multiagent Q-learning Dynamics with epsilon-greedy Exploration - 2010.pdf*
        + *Training with Exploration Improves a Greedy Stack LSTM Parser – 2015*
        + *Generalized Hindsight for Reinforcement Learning – 2020*
        + *Learning Multi-Level Hierarchies with Hindsight – 2019*
        + *Hierarchical Reinforcement Learning with Hindsight – 2018*
        + *Hindsight policy gradients - 2019*
  3. Optimização de recompensas
     + - *Deep Reinforcement Learning With Optimized Reward Functions for Robotic Trajectory Planning - xie2019*
       - *A Reward Optimization Method Based on Action Subrewards in Hierarchical Reinforcement Learning - fu2014*
       - *Rewards Prediction-Based Credit Assignment for Reinforcement Learning With Sparse Binary Rewards - seo2019*
  4. Séries Temporais
     1. Definição
        + *Time-series forecasting – 2001*
        + *Predicting chaotic time series - farmer1987*
     2. Metodologia para Coleta de dados
        + *Mining Time Series Data – 2005*
        + *Time-series data mining - esling2012*
        + *Searching and mining trillions of time series subsequences under dynamic time warping - rakthanmanon2012.pdf*
     3. Agricultural Commodities prices databases
        + Time Series Data Analysis on Agriculture Food Production
     4. Forecast
        + Forecast
          - Learning to Forecast Price - kelley2002.pdf
          - Short-Term Load Forecast of Microgrids by a New Bilevel Prediction Strategy - amjady2010
          - Reinforcement Learning Control for Water-Efficient Agricultural Irrigation - sun2017
          - Deep reinforcement learning for selecting demand forecast models to empower Industry 3.5 and an empirical study for a semiconductor component distributor – 2018
          - \_Crop Yield Prediction Using Deep Reinforcement Learning Model for Sustainable Agrarian Applications - elavarasan2020.pdf
     5. Taxonomia dos problemas de previsão de series temporais
        + Inputs vs. Outputs
        + Endogenous vs. Exogenous
        + Regression vs. Classification
        + Unstructured vs. Structured
        + Univariate vs. Multivariate
        + Single-step vs. Multi-step
        + Static vs. Dynamic
        + Contiguous vs. Discontiguous

1. **Metodologia**
2. **Resultados**
3. **Conclusão e trabalhos futuros**
   1. Conclusões
   2. Trabalhos futuros

**Ambiente de desenvolvimento**

Guia de Instalação

**Onde é aplicado?**

* *RL*
  + *Predicting Stock Market Trends Using Machine Learning and Deep Learning Algorithms Via Continuous and Binary Data; a Comparative Analysis – 2020*
  + *Applications of Deep Learning and Reinforcement Learning to Biological Data - mahmud2018.pdf*
  + *\_Deep Reinforcement Learning\_ An Overview – 2017*
  + *A review On reinforcement learning: Introduction and applications in industrial process control - nian2020.pdf*
* *RNN*
  + *Real-time recurrent learning neural network for stream-flow forecasting - chang2002.pdf*

Machine Learning

* Types of machine learning systems
  + Supervised
    - *A Comparative Study of Training Algorithms for Supervised Machine Learning – 2012*
    - *A review of supervised machine learning algorithms and their applications to ecological data - crisci2012*
  + Unsupervised
    - Unsupervised Learning - 1999
    - *Building High-level Features Using Large Scale Unsupervised Learning - 2012*
    - Unsupervised Machine Learning for Networking\_ Techniques, Applications and Research Challenges – 2019
  + Supervised/ Unsupervised
    - Comparing supervised and unsupervised category learning - Love2002.pdf
  + Semi-Supervised
    - Semi-Supervised and Unsupervised Extreme Learning Machines - gaohuang2014.pdf
    - Chapter 22 - Semi-Supervised Learning - zhou2014.pdf
    - MixMatch\_ A Holistic Approach to Semi-Supervised Learning - 2019
  + Reinforcement Learning
    - ***Use the references from the reinforcement learning chapter***
  + Online and Batch Learning
    - *On-Line Algorithms in Machine Learning - blum1998.pdf*
    - *Comparison of offline and real-time human activity recognition results using machine learning techniques - suto2018.pdf*
    - *Multi Kernel Learning with Online-Batch Optimization - orabona2012a.pdf*
  + Instance-based Learning
    - *Instance-based learning algorithms - aha1991.pdf*
    - *Combining instance-based learning and logistic regression for multilabel classification – 2009*
    - *Instance-Based Utile Distinctions for Reinforcement Learning with Hidden State - 1995*
  + Model-based Learning
    - ***Model-based and Model-free in reinforcement learning section***
* Neural Networks
  + - *https://www.dataversity.net/brief-history-deep-learning/*
    - *A logical calculus of the ideas immanent in nervous activity - mcculloch1943*
    - *Neural networks and physical systems with emergent collective computational abilities - 1982*
    - *On the Computational Power of Neural Nets - siegelmann1995.pdf*
    - Deep Learning
      * *Ivakhnenko, A. G. (1971). Polynomial Theory of Complex Systems - Possível primeira aparição Deep Learning*
      * *Deep learning in neural networks\_ An overview - schmidhuber2015.pdf*
  + Neurons: Biological and artificial
    - Artificial neural networks\_ a tutorial - jain1996.pdf
  + First mathematical modeling of neural network
    - *A Logical Calculus of the Ideas Immanent in Nervous Activity – 1943*
  + The Perceptron
    - *The perceptron - A probabilistic model for information storage and organization in the brain – 1958*
      * *Perceptrons\_ An introduction to computational geometry (1969, MIT Press) – 1969 ---- Não resolve XOR*
        + *In 1969, Minsky and Papert’s book Perceptrons was published. It was a harsh critique on Rosenblatt’s perceptrons. Minsky and Papert proved that perceptrons could only be trained to solve linear separable problems. For instance, one of the most dooming examples of non-linear separable problems is the exclusive OR (XOR).*
      * *Artificial intelligence- a general survey – Lighthill1973*
        + *The Lighthill report (published in 1973) was an evaluation of the current state of AI at that time written for the British Science Research Council. The report came to the conclusion that the promises of AI researchers were exaggerated: “in no part of the field have discoveries made so far produced the major impact that was then promised.”*
  + Feedforward
    - *Learning algorithms and probability distributions in feed-forward and feed-back networks – 1987*
    - *Training feedforward networks with the Marquardt algorithm - 1994*
  + Multilayer Perceptron Feedforward
    - *On the Capabilities of Multilayer Perceptrons – 1988*
    - *A General Multilayer Perceptrons Feed Forward Neural Network Algorithm for Learning Capability - 2013*
  + Multilayer Perceptron Backpropagation
    - Taylor expansion of the accumulated rounding error - 1970 - Backpropagation
      * Apesar de existir era ineficiente devido o poder computacional da época, pelo menos até 1985
    - *Learning representations by back-propagating errors - 1986 Mostra que BackPropation pode ser usado para obter melhores resultados.pdf*
    - *A theoretical framework for Back-Propagation – 1988*
    - *Improving the Convergence of Back-Propagation Learning with Second-Order Methods – 1989*
    - *Theory of the backpropagation neural network – 1989*
    - *Back-propagation fails to separate – 1989*
    - *On the problem of local minima in backpropagation – 1992*
    - *Efficient BackProp*
      * *On derivation of MLP backpropagation from the Kelley-Bryson optimal-control gradient formula and its application*
  + Gradiente Estocástico
  + *Support - Vector Networks – 1995*
  + Activation Functions
    - https://missinglink.ai/guides/neural-network-concepts/7-types-neural-network-activation-functions-right/
    - Binary Step Function
    - Linear Activation Function
      * *Learning Activation Functions to Improve Deep Neural Networks*
    - Non-Linear Activation Functions
      * + *Learning Activation Functions to Improve Deep Neural Networks*
      * Sigmoid / Logistic
        + Performance Analysis of Various Activation Functions in Generalized MLP Architectures of Neural Networks – 2011
        + Implementation of a Sigmoid Activation function for Neural Network - 2012
        + Analysis of different activation functions using back propagation neural networks – 2013
      * TanH / Hyperbolic Tangent
        + Why Tanh\_ Choosing a Sigmoidal Function - 1992
      * ReLU (Rectified Linear Unit)
        + Parameterised Sigmoid and ReLU Hidden Activation Functions for DNN Acoustic Modelling - 2015
        + Deep Learning using Rectified Linear Units (ReLU) – 2018
        + Searching for Activation Functions – 2017
        + Research on convolutional neural network based on improved Relu piecewise activation function - lin2018
        + Leaky ReLU

Comparative Study of Convolution Neural Network’s Relu and Leaky-Relu Activation Functions - 2019

* + - * + Parametric ReLU

Alcoholism identification via convolutional neural network based on parametric ReLU, dropout, and batch normalization - wang2018

* + - * *Softmax*
        + *On The Pairing Of The Softmax Activation And Cross{Entropy Penalty Functions And The Derivation Of The Softmax Activation Function - 1997*
        + *Simplified Hardware Implementation of the Softmax Activation Function - kouretas2019.pdf*
        + *ACTIVATION FUNCTIONS IN NEURAL NETWORKS - 2020*
      * Swish
        + *Swish: a Self-Gated Activation Function*
        + *Flatten-T Swish\_ a thresholded ReLU-Swish-like activation function for deep learning - 2018*
  + Optimizers
    - Gradient Descent variants
      * Batch gradient descent
        + *The general inefficiency of batch training for gradient descent learning*
      * Stochastic gradient descent
        + *A Stochastic Approximation Method - SGD First Paper*
        + *Stochastic Estimation of the Maximum of a Regression Function - SGD 2*
        + *An overview of gradient descent optimization*
      * Mini-batch gradient descent
        + *Better Mini-Batch Algorithms via Accelerated Gradient Methods*
    - Momentum
      * *Some methods of speeding up the convergence of iteration methods - Momentum*
      * *On the importance of initialization and momentum in deep learning*
    - Grandiente acelerado Nesterov
      * *A method for solving the convex programming problem with convergence rate - Nesterov Original Paper*
      * *Accelerated Distributed Nesterov Gradient Descent*
    - AdaGrad
      * *Adaptive Subgradient Methods for Online Learning and Stochastic Optimization -ADAGRAD*
    - Adadelta
      * *ADADELTA: An Adaptive Learning Rate Method*
    - RMSProp
      * *Neural Networks for Machine - Learning Lecture 6a Overview of mini-batch gradient descent - RMSProp*
    - Adam
      * *ADAM: A METHOD FOR STOCHASTIC OPTIMIZATION*
    - ***To summarize, RMSProp, AdaDelta and Adam are very similar algorithm and since Adam was found to slightly outperform RMSProp, Adam is generally chosen as the best overall choice.***
  + Overfitting e Underfitting
    - *Machine Learning Basics - Overfitting e Underfitting – Slide*
    - *Process mining\_ a two-step approach to balance between underfitting and overfitting - 2010*
    - *Methods To Avoid Over-fitting And Under-fitting In Supervised Machine Learning (comparative Study) – 2015*
    - *A pruning based method to learn both weights and connections for LSTM – 2015*
    - *Overfitting and Underfitting Analysis for Deep Learning Based End-to-end Communication Systems - 2019*
    - *Study of Dependency on number of LSTM units for Character based Text Generation models - chakraborty2020.pdf*
    - Early-Stopping
      * Early Stopping - But When - prechelt1998.pdf
      * Training Recurrent Answering Units with Joint Loss Minimization for VQA - 2016
      * HINDSIGHT\_ An R-Based Framework Towards Long Short Term Memory (LSTM) Optimization – 2018Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems - Aurélien Géron – 2019
      * https://keras.io/api/callbacks/early\_stopping/
  + Regularization
    - *Networks for Approximation and Learning - regularization – 1990*
    - *Feature selection, L1 vs. L2 regularization, and rotational invariance - ng2004*
    - *On Manifold Regularization - TR-2004-05*
    - *Recurrent Neural Network Regularization - 2015*
    - *Regularizing and Optimizing LSTM Language Models – 2017*
    - *What is the Effect of Importance Weighting in Deep Learning - 2019*
    - Dropout
      * *Dropout Training as Adaptive Regularization – 2013*
      * *The Implicit and Explicit Regularization Effects of Dropout – 2020*
      * *Adversarial Dropout Regularization - 2017*
    - Data Augmentation
      * The Art of Data Augmentation - vandyk2001
      * Understanding Data Augmentation for Classification\_ When to Warp - wong2016
      * Data Augmentation for Recognition of Handwritten Words and Lines Using a CNN-LSTM Network - wigington2017.pdf
      * Data Augmentation and Dense-LSTM for Human Activity Recognition using WiFi Signal - zhang2020
  + *ALEXNET*
    - *ImageNet Classification with Deep Convolutional Neural Networks -* ***Alexnet*** *– 2012*

CNN (Convolutional Neural Network)

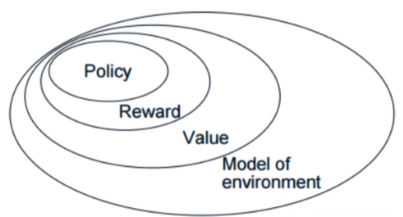
* + *Cognitron - A self-organizing multilayered neural network - First "CNN" 1975*
  + Exploring Convolutional Neural Network Structures and Optimization Techniques for Speech Recognition - cnn\_2013
  + *Fast R-CNN - 2015.pdf*
* Convolutional Layer
  + *Going deeper with convolutions - 2014*
  + *Understanding of a convolutional neural network - albawi2017*
  + Filtros
    - Non-Linear Convolution Filters for CNN-Based Learning - 2017
    - Learning Structure and Strength of CNN Filters for Small Sample Size Training – 2018
    - Learning CNN Filters From User-Drawn Image Markers for Coconut-Tree Image Classification - desouza2020.pdf
  + Feature Maps
    - A CNN-based face detector with a simple feature map and a coarse-to-fine classifier - Withdrawn - chen2015.pdf
    - Object Detection Networks on Convolutional Feature Maps - ren2016.pdf
* Batch Normalization
  + Batch Normalization\_ Accelerating Deep Network Training by Reducing Internal Covariate Shift - ioffe2015.pdf
  + **Batch normalized** recurrent neural networks - laurent2016.pdf
  + How Does Batch Normalization Help Optimization - 2019
* Pooling Layer
  + Striving for Simplicity\_ The All Convolutional Net – 2014
  + Deep Residual Learning for Image Recognition – 2015
  + Temporal Pyramid Pooling-Based Convolutional Neural Network for Action Recognition - wang2016.pdf
* Fully Connected Layer
  + Convolutional, Long Short-Term Memory, fully connected Deep Neural Networks – 2015
  + Impact of fully connected layers on performance of convolutional neural networks for image classification - basha2019.pdf

Redes Neurais Recorrentes (RNN)

* Structure
  + *An Empirical Exploration of Recurrent Network Architectures - 2015*
  + *Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network - sherstinsky2020*
  + *Visualizing and Understanding Recurrent Networks – 2015*
  + *Batch normalized recurrent neural networks - laurent2016*
  + *Hidden State*
    - *Instance-Based Utile Distinctions for Reinforcement Learning with Hidden State – 1995*
    - *Visual Analysis of Hidden State Dynamics in Recurrent Neural Networks – 2016*
* Vanishing/Exploding Gradient
  + *Learning long-term dependencies with gradient descent is difficult - bengio1994*
  + *Overcoming the vanishing gradient problem in plain recurrent networks*
  + *Gradient Flow in Recurrent Nets: the Difficulty of Learning Long-Term Dependencies - Vanish Gradient Problem 2001 .pdf*
  + *On the difficulty of training Recurrent Neural Networks - Gradient Problems – 2012*
    - *important*
  + Exploding Gradient
    - Truncated Backpropagation solves the problem of exploding gradient, but generate another problem, that is the weights that won’t be completely updated
      * You can use penalties to
    - Gradient Clipping
    - RMSProp – divide a taxa de aprendizagem por uma média exponencialmente decrescente do quadrado do gradiente
  + Vanishing Gradient
    - *Backpropagation Through Time*
    - *Weight initialization (Xavier Initialization)*
    - *Echo State Networks*
    - *LSTMs*
* LTSM (Long Short-Term Memory) Network
  + [*http://colah.github.io/posts/2015-08-Understanding-LSTMs/*](http://colah.github.io/posts/2015-08-Understanding-LSTMs/)
  + *https://medium.com/mlreview/understanding-lstm-and-its-diagrams-37e2f46f1714*
  + *lstm – 1997*
  + Structure
    - Input Gate
      * Add memory
    - Forget Gate
      * Free memory
    - Output gate
      * Read memory
  + Memory cells
    - *Investigating recurrent neural network memory structures using neuro-evolution - ororbia2019*
    - *Do RNN and LSTM have Long Memory – 2020*
  + Activation Function
    - Sigmoid
    - Tanh
  + *Gradient Flow in Recurrent Nets: the Difficulty of Learning Long-Term Dependencies - Vanish Gradient Problem 2001 .pdf*
  + *long short-term memory (LSTM) network of 2003 by Hochreiter \_ Schmidhuber*
  + *Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling - 2014*
  + *Using ant colony optimization to optimize long short-term memory recurrent neural networks - elsaid2018*
  + *A Hybrid CNN–LSTM Network for the Classification of Human Activities Based on Micro-Doppler Radar - zhu2020*
  + LSTM\_ A Search Space Odyssey – 2015
    - Different types of LSTMs
  + GRU
    - *Gate-variants of Gated Recurrent Unit (GRU) neural networks - dey2017.pdf*
    - *Recurrent Neural Networks for Multivariate Time Series with Missing Values – 2018*
* Backpropagation
  + *Reinforcement learning by backpropagation through an LSTM model-critic – 2007*
  + *A Gentle Tutorial of Recurrent Neural Network with Error Backpropagation – 2016*
  + *Comparison of Predictive Algorithms- Backpropagation, SVM, LSTM and Kalman Filter for Stock Market - 2019*
  + *Dense Recurrent Neural Networks for Accelerated MRI- History-Cognizant Unrolling of Optimization Algorithms - hosseini2020*
* Backpropagation Through Time
  + Recurrent Neural Networks Trained With Backpropagation Through Time Algorithm to Estimate Nonlinear Load Harmonic Currents - mazumdar2008
* Truncated Backpropagation Through Time
  + *Unbiasing Truncated Backpropagation Through Time – 2017*
  + *On Training Recurrent Networks with Truncated Backpropagation Through time in Speech Recognition - tang2018*
  + *Adaptively Truncating Backpropagation Through Time to Control Gradient Bias - 2020*
* Regularization and Optimizing
  + *Regularizing and Optimizing LSTM Language Models – 2017*
  + <https://en.wikipedia.org/wiki/Vanishing_gradient_problem#Long_short-term_memory>

Reinforcement Learning

* [*http://incompleteideas.net/book/first/ebook/node12.html*](http://incompleteideas.net/book/first/ebook/node12.html)
* Foundations
  + *An Example of Statistical Investigation of the Text Eugene Onegin Concerning the Connection of Samples in Chains - Markov Chains 1913 - Translate 2007*
  + *A Mathematical Theory of Communication – 1948*
  + *Finding Structure in Reinforcement Learning - NIPS-1994*
  + *Learning to predict by the methods of temporal difference - Sutton1988.pdf*
  + *Learning from delayed reward - 1989.pdf*
  + *Reinforcement Learning for Robots Using Neural Networks - 1993*
  + *Reinforcement Learning and the Reward Engineering Principle - 2014*
  + *From Reinforcement Learning to Optimal Control: A unified framework for sequential decisions - 2019.pdf*
  + *Holistic Reinforcement Learning - The Role of Structure and Attention - radulescu2019*
  + *Deep Reinforcement Learning for Dexterous Manipulation with Concept Networks – 2017*
  + **State, Action, Reward, Discount** – Concepts of Dynamic programming
* The Bellman equation
  + - *On the role of dynamic programming in statistical communication theory - bellman 1957*
    - *Dynamic programming and stochastic control processes - bellman1958* 
      * ***introduced the discrete stochastic version of the optimal control problem known as Markovian decision processes (MDPs)\*\*\****
    - *On adaptive control processes - bellman1959.pdf*
    - *Applied dynamic programming – 1962*
* *Planning*
  + *Integrated Architectures for Learning, Planning, and Reacting Based on Approximating Dynamic Programming - sutton1990.pdf*



* Markov Process
* Markov Decision Process
  + *Dynamic Programming - rhoward1960*
    - ***devised the policy iteration method for MDPs. All of these are essential elements underlying the theory and algorithms of modern reinforcement learning.***
  + *Means and variances of time averages in Markovian environments - 1987.pdf*
  + *Numerical Transient Analysis Of Markov Models – 1988*
  + *Transient analysis of cumulative measures of markov model behavior – 1989*
  + *Hierarchical Control and Learning for Markov Decision Processes - 1998 – 1990*
  + *A survey of algorithmic methods for partially observed Markov decision processes - lovejoy1991*
  + *A Survey of Applications of Markov Decision Processes - 1991*
  + *Markov and the Birth of Chain Dependence Theory – 1996*
  + *Markov Decision Processes\_ Concepts and Algorithms – 2009 \*\*\**
  + Markov reward process
    - *Markov and Markov reward model transient analysis - An overview of numerical approaches - reibman1989*
    - *Performability Analysis Using Semi-Markov Reward Processes – 1990*
    - *Automated Generation And Analysis Of Markov Reward Models Using Stochastic Reward Nets - 1993*
    - *Markov Reward Models and Markov Decision Processes in Discrete and Continuous Time-Performance Evaluation and Optimization – gouberman2014*
    - **Living Penalty**
* Policy
  + *A policy-blending formalism for shared control - dragan2013*
  + *Policy Gradient Methods for Reinforcement Learning with Function Approximation – 1999*
  + *Composite Task-Completion Dialogue Policy Learning via Hierarchical Deep Reinforcement Learning – 2017*
  + Fast Policy Learning through Imitation and Reinforcement – 2018
  + *Improving RTS Game AI by Supervised Policy Learning, Tactical Search, and Deep Reinforcement Learning - barriga2019.pdf*
  + **Policy Gradients**
    - *Reinforcement Learning Through Gradient Descent – 1999*
    - *Gradient Descent for General Reinforcement Learning – 1999*
    - *Policy gradient methods for reinforcement learning with function approximation - 2000*
    - *Policy Gradient Reinforcement Learning for Fast Quadrupedal Locomotion - kohl2004*
    - *Restricted gradient-descent algorithm for value-function approximation in reinforcement learning – 2007*
    - *Restricted gradient-descent algorithm for value-function approximation in reinforcement learning - damottasallesbarreto2008*
    - *The Optimal Reward Baseline for Gradient·Based Reinforcement Learning – 201*
* *Model-based and model-free learning*
  + *Model-based*
    - *Model-based learning for mobile robot navigation from the dynamical systems perspective - tani1996*
    - *Introduction to model-based teaching and learning in science education - gobert2000*
    - *Multiple Model-Based Reinforcement Learning - doya2002*
    - *Model-Based Reinforcement Learning for Atari - 2020*
  + *Model-Free Learning*
    - *Model-free Q-learning designs for linear discrete-time zero-sum games with application to H-infinity control - al-tamimi2007*
    - *Learning model-free robot control by a Monte Carlo EM algorithm - vlassis2009*
    - *Generalized TD Learning - 2011*
    - *Model-Free reinforcement learning with continuous action in practice - degris2012*
    - *Neural Computations Underlying Arbitration between Model-Based and Model-free Learning - lee2014*
    - *Model-Based Value Estimation for Efficient Model-Free Reinforcement Learning – 2018*
* *Monte Carlo? RBF Network?*
* *TD*
* *Q-Learning*
  + *Quality of certain action in a given state*
  + *Watkins-Dayan1992\_Article\_Q-learning*
  + Time difference learning
    - *Learning to predict by the methods of temporal differences – 1988*
      * *Classic refference*
    - *Temporal Difference Learning and TD-Gammon – 1995*
    - *Incremental Least-Squares Temporal Difference Learning - 2006*
    - Políticas de Extrapolação
      * Q-learning aproximado e Deep Q-Learning
* *Deep Q-Learning*
  + *Experiência de Replay*
    - *Prioritized Experience Replay – 2015*
    - *Distributed Prioritized Experience Replay – 2018*
    - *Experience Replay Optimization – 2019*
* *Forward model*
  + *Self-improving reactive agents based on reinforcement learning, planning and teaching – 1992*
  + *Ants and Reinforcement Learning\_ A Case Study in Routing in Dynamic Networks – 1998*
  + *Forward-Backward Reinforcement Learning – 2018*
* Backpropagation
  + *On the use of backpropagation in associative reinforcement learning - williams1988.pdf*
  + *Input Generalization in Delayed Reinforcement Learning\_ An Algorithm and Performance Comparisons – 1991*
  + *Reinforcement learning by backpropagation through an LSTM model-critic – 2007*
* *Desafios*
  + *Introduction- The Challenge of Reinforcement Learning – 1992*
  + *Learning agents for uncertain environments (extended abstract) – 1998*
  + *Constrained Markov decision processes – 1999*
  + *Challenges of Real-World Reinforcement Learning – 2019*
  + *An empirical investigation of the challenges of real-world reinforcement learning – 2020*
* *Credit Assignment Problem (Caused by sparce reward setting)*
  + *TEMPORAL CREDIT ASSIGNMENT IN REINFORCEMENT LEARNING - 1984.pdf*
  + *Unifying Temporal and Structural Credit Assignment Problems - 2004.pdf*
  + *Solving the credit assignment problem: explicit and implicit learning of action sequences with probabilistic outcomes - fu2007.pdf*
* Sparce rewards – Really Hard and Not An Optimal Solution
  + *Reinforcement Learning with Unsupervised Auxiliary Tasks - 2016*
  + *Leveraging Demonstrations for Deep Reinforcement – 2017*
  + *Overcoming Exploration in Reinforcement Learning with Demonstrations - nair2018*
  + *Hindsight Experience Replay - 2018*
  + *Deep-Reinforcement-Learning-Based Autonomous UAV Navigation With Sparse Rewards - wang2020*
* Reward Shape – Not Scalable and Suffer with the alignment problem
  + Reward shaping for reinforcement learning by emotion expressions - hwang2014
* The Alignment Problem
  + The Alignment Problem for Bayesian History-Based Reinforcement Learners - everitt2018.pdf
  + Using deep reinforcement learning approach for solving the multiple sequence alignment problem – 2019
* Exploration vs Exploitation dilemma
  + Exploration and Exploitation in Organizational Learning - March\_1991.pdf
  + The neuroscientific foundations of the exploration−exploitation dilemma - 2010
  + Reinforcement learning\_ exploration–exploitation dilemma in multi-agent foraging task - yogeswaran2012.pdf
  + The Exploration-Exploitation Dilemma\_ A Multidisciplinary Framework - berger-tal2014.pdf
  + MULEX\_ Disentangling Exploitation from Exploration in Deep RL – 2017
  + Solution *epsilon* Gradient
  + Curiosity Driven Exploration
    - *Agent can stuck in local minimum with small, recurring rewards*
    - *Curiosity-driven Exploration by Self-supervised Prediction – 2017*
    - *Adaptive ε-Greedy Exploration in Reinforcement Learning Based on Value Differences – 2010 \*\*\**
    - *Classes of Multiagent Q-learning Dynamics with epsilon-greedy Exploration - 2010.pdf*
    - *Training with Exploration Improves a Greedy Stack LSTM Parser – 2015*
    - *Generalized Hindsight for Reinforcement Learning – 2020*
    - *Learning Multi-Level Hierarchies with Hindsight – 2019*
    - *Hierarchical Reinforcement Learning with Hindsight – 2018*
    - *Hindsight policy gradients - 2019*
* Optimize Rewards
  + *Deep Reinforcement Learning With Optimized Reward Functions for Robotic Trajectory Planning - xie2019*
  + *A Reward Optimization Method Based on Action Subrewards in Hierarchical Reinforcement Learning - fu2014*
  + *Rewards Prediction-Based Credit Assignment for Reinforcement Learning With Sparse Binary Rewards - seo2019*