1. [https://theconversation.com/what-was-the-first-computer-122164#:~:text=The%20first%20mechanical%20computer%2C%20The,computer%20we%20all%20use%20today.&text=The%20ABC%20weighed%20over%20700,had%20small%20capacitors%20on%20it](https://theconversation.com/what-was-the-first-computer-122164" \l ":~:text=The%20first%20mechanical%20computer%2C%20The,computer%20we%20all%20use%20today.&text=The%20ABC%20weighed%20over%20700,had%20small%20capacitors%20on%20it).
2. Das, K. and Behera, R.N., 2017. A survey on machine learning: concept, algorithms and applications. *International Journal of Innovative Research in Computer and Communication Engineering*, *5*(2), pp.1301-1309.
3. Hebb, D.O., 2005. *The organization of behavior: A neuropsychological theory*. Psychology Press.
4. Turing, Alan M. "Computing machinery and intelligence." In *Parsing the turing test*, pp. 23-65. Springer, Dordrecht, 2009.
5. <https://www.smdi.com/the-evolution-of-machine-learning/>
6. McCarthy, J., Minsky, M.L., Rochester, N. and Shannon, C.E., 2006. A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. *AI magazine*, *27*(4), pp.12-12.
7. Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. Psychological Review, 65(6), 386–408.
8. T. Cover and P. Hart, "Nearest neighbor pattern classification," in IEEE Transactions on Information Theory, vol. 13, no. 1, pp. 21-27, January 1967
9. Rumelhart, D., Hinton, G. & Williams, R. Learning representations by back-propagating errors. *Nature* **323,**533–536 (1986).
10. Y. LeCun et al., "Backpropagation Applied to Handwritten Zip Code Recognition," in Neural Computation, vol. 1, no. 4, pp. 541-551, Dec. 1989.
11. Sah, S., 2020. Machine learning: a review of learning types.
12. Sarker, I.H. Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN COMPUT. SCI.* **2,**160 (2021)
13. Nasteski, V., 2017. An overview of the supervised machine learning methods. *Horizons. b*, *4*, pp.51-62.
14. Sandhya N. dhage, Charanjeet Kaur Raina. (2016) *A review on Machine Learning Techniques*. In International Journal on Recent and Innovation Trends in Computing and Communication, Volume 4 Issue 3.
15. Osisanwo, F.Y., Akinsola, J.E.T., Awodele, O., Hinmikaiye, J.O., Olakanmi, O. and Akinjobi, J., 2017. Supervised machine learning algorithms: classification and comparison. *International Journal of Computer Trends and Technology (IJCTT)*, *48*(3), pp.128-138.
16. Mladenić, Dunja, Janez Brank, Marko Grobelnik, and Natasa Milic-Frayling. "Feature selection using linear classifier weights: interaction with classification models." In *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 234-241. 2004.
17. Rish, Irina. "An empirical study of the naive Bayes classifier." In *IJCAI 2001 workshop on empirical methods in artificial intelligence*, vol. 3, no. 22, pp. 41-46. 2001.
18. Kleinbaum, D.G., Dietz, K., Gail, M., Klein, M. and Klein, M., 2002. *Logistic regression* (p. 536). New York: Springer-Verlag.
19. Rosenblatt, Frank. *Principles of neurodynamics. perceptrons and the theory of brain mechanisms*. Cornell Aeronautical Lab Inc Buffalo NY, 1961.
20. Baum, E.B., 1988. On the capabilities of multilayer perceptrons. *Journal of complexity*, *4*(3), pp.193-215.
21. Pisner, Derek A., and David M. Schnyer. "Support vector machine." In *Machine learning*, pp. 101-121. Academic Press, 2020.
22. Kotsiantis, S.B., 2013. Decision trees: a recent overview. *Artificial Intelligence Review*, *39*(4), pp.261-283.
23. Biau, G. and Scornet, E., 2016. A random forest guided tour. *Test*, *25*(2), pp.197-227.
24. Jain, A.K., Mao, J. and Mohiuddin, K.M., 1996. Artificial neural networks: A tutorial. *Computer*, *29*(3), pp.31-44.
25. Friedman, N., Geiger, D. and Goldszmidt, M., 1997. Bayesian network classifiers. *Machine learning*, *29*(2), pp.131-163.
26. Osisanwo, F.Y., Akinsola, J.E.T., Awodele, O., Hinmikaiye, J.O., Olakanmi, O. and Akinjobi, J., 2017. Supervised machine learning algorithms: classification and comparison. *International Journal of Computer Trends and Technology (IJCTT)*, *48*(3), pp.128-138.
27. Celebi, M. Emre, and Kemal Aydin, eds. *Unsupervised learning algorithms*. Berlin: Springer International Publishing, 2016.
28. Zhang, J., Chen, W., Gao, M. and Shen, G., 2017. K-means-clustering-based fiber nonlinearity equalization techniques for 64-QAM coherent optical communication system. *Optics express*, *25*(22), pp.27570-27580.
29. Guo, Gongde, Hui Wang, David Bell, Yaxin Bi, and Kieran Greer. "KNN model-based approach in classification." In *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"*, pp. 986-996. Springer, Berlin, Heidelberg, 2003.
30. Murtagh, F. and Contreras, P., 2012. Algorithms for hierarchical clustering: an overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *2*(1), pp.86-97.
31. Abdi, H. and Williams, L.J., 2010. Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, *2*(4), pp.433-459.
32. Baker, K., 2005. Singular value decomposition tutorial. *The Ohio State University*, *24*.
33. Stone, James V. "Independent component analysis: an introduction." *Trends in cognitive sciences* 6, no. 2 (2002): 59-64.
34. Atallah, D.M., Badawy, M., El-Sayed, A. and Ghoneim, M.A., 2019. Predicting kidney transplantation outcome based on hybrid feature selection and KNN classifier. *Multimedia Tools and Applications*, *78*(14), pp.20383-20407.
35. Arulkumaran, K., Deisenroth, M.P., Brundage, M. and Bharath, A.A., 2017. A brief survey of deep reinforcement learning. *arXiv preprint arXiv:1708.05866*.
36. Müller, Berndt, Joachim Reinhardt, and Michael T. Strickland. *Neural networks: an introduction*. Springer Science & Business Media, 1995.
37. Müller, Berndt, Joachim Reinhardt, and Michael T. Strickland. *Neural networks: an introduction*. Springer Science & Business Media, 1995.
38. McCulloch, W.S. and Pitts, W., 1943. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, *5*(4), pp.115-133.
39. <https://towardsdatascience.com/rosenblatts-perceptron-the-very-first-neural-network-37a3ec09038a>
40. Widrow, Bernard, and Marcian E. Hoff. *Adaptive switching circuits*. Stanford Univ Ca Stanford Electronics Labs, 1960.
41. Minsky, M. and Papert, S., 1969. Perceptrons.
42. Werbos, Paul. (1974). Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Science. Thesis (Ph. D.). Appl. Math. Harvard University.
43. Linnainmaa, S., 1976. Taylor expansion of the accumulated rounding error. *BIT Numerical Mathematics*, *16*(2), pp.146-160.
44. Wythoff, B.J., 1993. Backpropagation neural networks: a tutorial. *Chemometrics and Intelligent Laboratory Systems*, *18*(2), pp.115-155.
45. Kunihiko Fukushima (1980). *Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. , 36(4), 193–202.*
46. LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W. and Jackel, L.D., 1989. Backpropagation applied to handwritten zip code recognition. *Neural computation*, *1*(4), pp.541-551.
47. LeCun, Y., 1989. Generalization and network design strategies. *Connectionism in perspective*, *19*(143-155), p.18.
48. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M. and Berg, A.C., 2015. Imagenet large scale visual recognition challenge. *International journal of computer vision*, *115*(3), pp.211-252.
49. Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, *25*.
50. Nair, Vinod, and Geoffrey E. Hinton. "Rectified linear units improve restricted boltzmann machines." In *Icml*. 2010.
51. Zeiler, Matthew D., and Rob Fergus. "Visualizing and understanding convolutional networks." In *European conference on computer vision*, pp. 818-833. Springer, Cham, 2014.
52. Szegedy, Christian, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. "Going deeper with convolutions." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1-9. 2015.
53. Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556* (2014).
54. He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770-778. 2016.
55. Srivastava, Rupesh Kumar, Klaus Greff, and Jürgen Schmidhuber. "Highway networks." *arXiv preprint arXiv:1505.00387* (2015).
56. He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Identity mappings in deep residual networks." In *European conference on computer vision*, pp. 630-645. Springer, Cham, 2016.
57. P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” in Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, vol. 1. IEEE, 2001, pp. 1–1
58. Zou, Zhengxia, Zhenwei Shi, Yuhong Guo, and Jieping Ye. "Object detection in 20 years: A survey." *arXiv preprint arXiv:1905.05055* (2019).
59. N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, vol. 1. IEEE, 2005, pp. 886–893
60. P. Felzenszwalb, D. McAllester, and D. Ramanan, “A discriminatively trained, multiscale, deformable part model,” in Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on. IEEE, 2008, pp. 1–8
61. Sermanet, Pierre, David Eigen, Xiang Zhang, Michaël Mathieu, Rob Fergus, and Yann LeCun. "Overfeat: Integrated recognition, localization and detection using convolutional networks." *arXiv preprint arXiv:1312.6229* (2013).
62. R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 580– 587.
63. K. E. Van de Sande, J. R. Uijlings, T. Gevers, and A. W. Smeulders, “Segmentation as selective search for object recognition,” in Computer Vision (ICCV), 2011 IEEE International Conference on. IEEE, 2011, pp. 1879–1886.
64. K. He, X. Zhang, S. Ren, and J. Sun, “Spatial pyramid pooling in deep convolutional networks for visual recognition,” in European conference on computer vision. Springer, 2014, pp. 346–361.
65. R. Girshick, “Fast r-cnn,” in Proceedings of the IEEE international conference on computer vision, 2015, pp. 1440–1448.
66. Ren, Shaoqing, Kaiming He, Ross Girshick, and Jian Sun. "Faster r-cnn: Towards real-time object detection with region proposal networks." *Advances in neural information processing systems* 28 (2015).
67. J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In CVPR, 2015.
68. Everingham, M., Van Gool, L., Williams, C.K.I. *et al.* The PASCAL Visual Object Classes (VOC) Challenge. *Int J Comput Vis* **88,**303–338 (2010).
69. Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *The Psychological Review: Monograph Supplements*, 2(4), i–109.
70. Thorndike, Edward L. “The Law of Effect.” *The American Journal of Psychology*, vol. 39, no. 1/4, 1927, pp. 212–22. *JSTOR*.
71. Sutton, Richard S., and Andrew G. Barto. *Reinforcement learning: An introduction*. MIT press, 2018.
72. Bellman, R., 1954. The theory of dynamic programming. *Bulletin of the American Mathematical Society*, *60*(6), pp.503-515.
73. Bellman, R., 1957. A Markovian decision process. *Journal of mathematics and mechanics*, pp.679-684.
74. Howard, R. A. (1960). Dynamic programming and Markov processes. John Wiley.
75. Arthur, S., 1959. Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, *3*(3), pp.210-229.
76. Minsky, M., 1961. Steps toward artificial intelligence. *Proceedings of the IRE*, *49*(1), pp.8-30.
77. Minsky, Marvin Lee. *Theory of neural-analog reinforcement systems and its application to the brain-model problem*. Princeton University, 1954.
78. Watkins, Christopher John Cornish Hellaby. "Learning from delayed rewards." (1989). PhD Thesis.