Using Neural Network to Predict Student Performance

Laly Xiong

Computer Science Senior Seminar

April 18, 2020

Overview

- **Introduction**
- **Background**
- **Method**
- **Results**
- **Conclusion**

Introduction

- Terms that are used consistently throughout presentation
  - **Neural Network**
  - **Data**
- Increased accessibility to student data over the years
- Predicting student performance can enhance students learning experience
- Help professors figure out who needs assistance

Overview

- Introduction
- **Background**
- Method
- Results
- Conclusion

Background

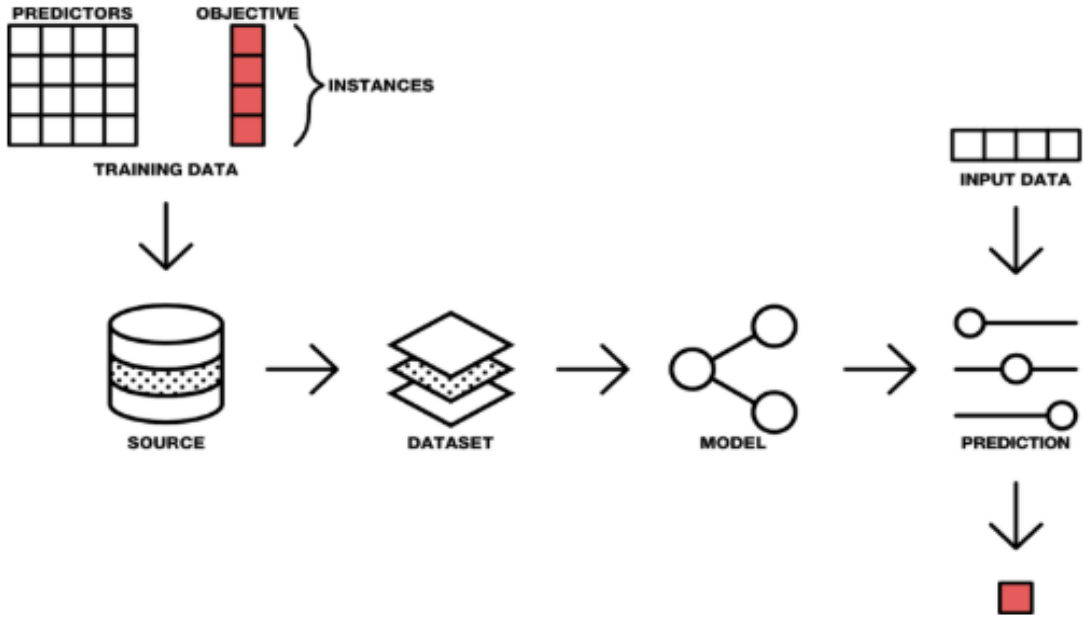
- **Machine Learning**
  - **Predictors**
  - **Classifiers**
  - **Confusion Matrices**
  - **Accuracy**
- Neural Network
- Training
- Decision Tree
- Educational Data-mining

Machine Learning

- Give computer ability to learn without being specifically programmed to do so
- Take in data then learn from that data to make predictions (training)
- Functions without some human assistance
- Examples such as speech recognition and prediction

Machine Learning - Predictors

- Variables in the data that can be used to predict the outcome



Machine Learning - Classifiers

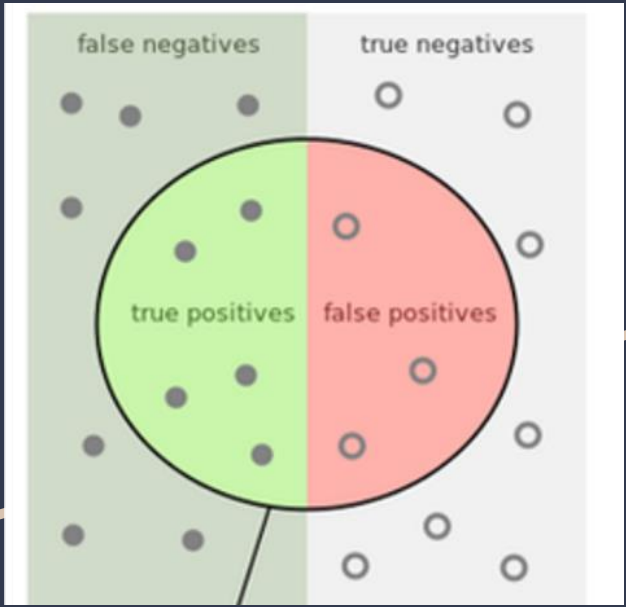
- Function or value that is used to assign labels to data points
- Labels can be the targeted output
- Predictive modeling used data and statistics to predict an outcome

Machine Learning - Confusion Matrices

- Table that shows performance of algorithm

		Predicted Class		
		Class 0	Class 1	Class 2
Actual Class	Class 0	0.80	0.15	0.05
	Class 1	0.10	0.76	0.14
	Class 2	0.03	0.11	0.86

Machine Learning - Accuracy



- One way to evaluate classification models

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

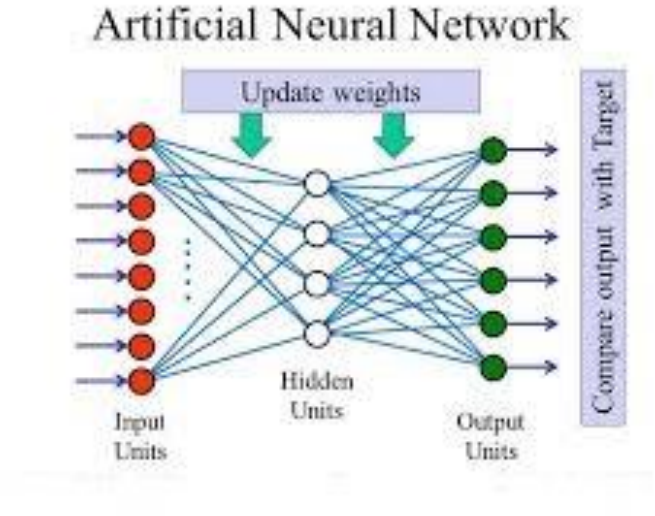
Where TP = True Positive, TN = True Negatives, FP = False Positives and FN = False Negatives.

Background

- Machine Learning
- **Neural Network**
  - **Layers**
  - **Edges**
  - **Weights**
  - **Bias**
  - **Activation Functions**
- Training
- Decision Tree
- Educational Data-mining

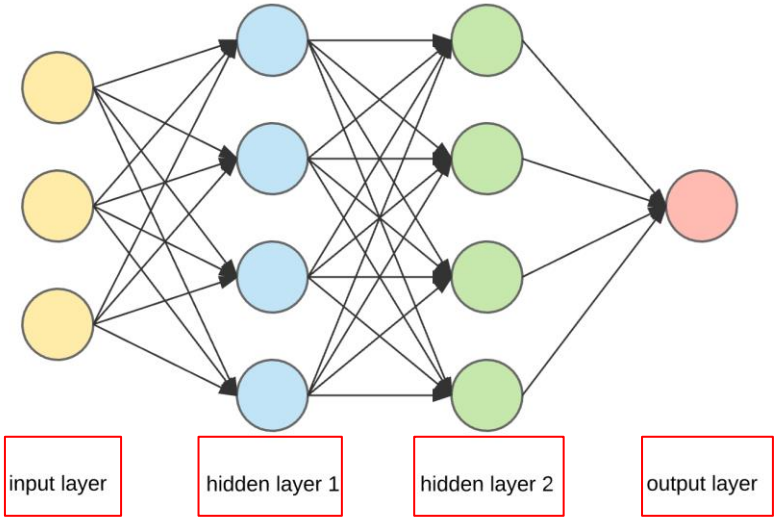
Neural Network

- Set of algorithms that are designed to learn from a data-set and recognize patterns
- Created to imitate firing of neurons in the brain
- Also called Artificial Neural Network (ANN)



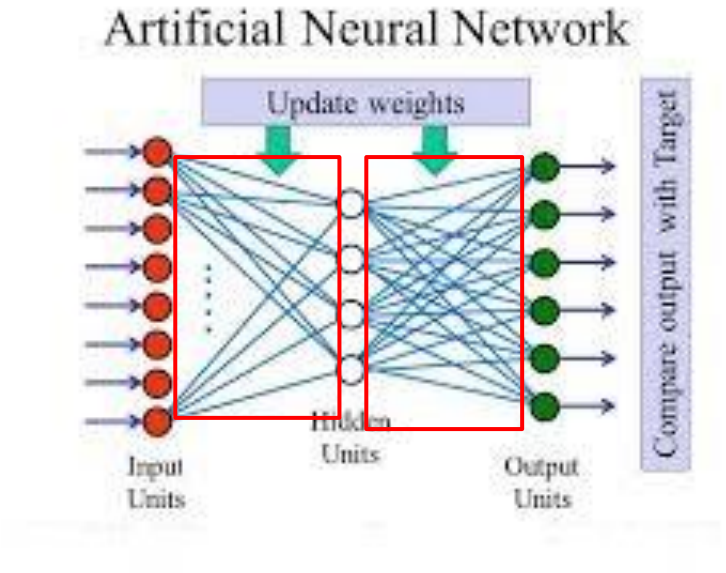
Neural Network - Layers

- A collection of nodes that interact with each other
- Has 3 layers:
  - Input- Where the data is inserted
  - Hidden- Where activation functions are performed
  - Output- Adjusted data

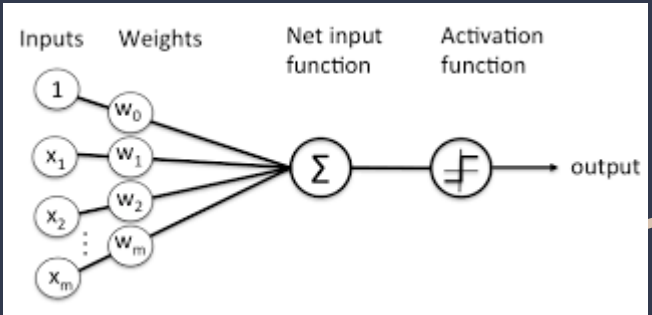


Neural Network - Edges

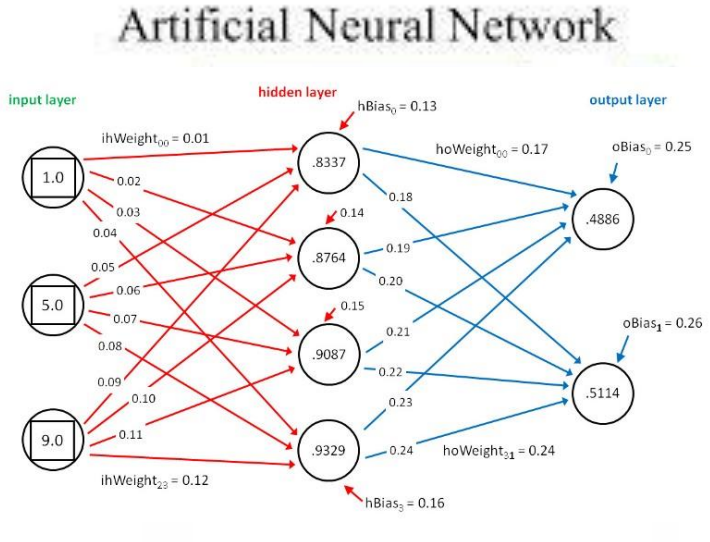
- Represent the connection between each node in the layers



Neural Network - Weights

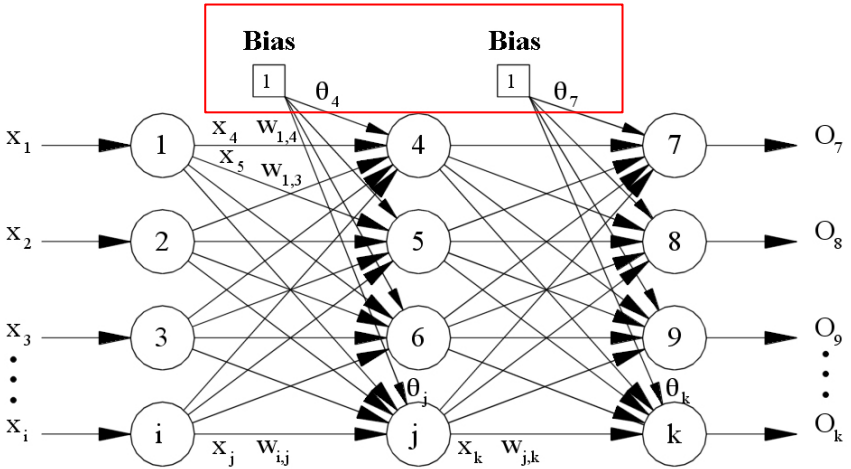


- Show the importance of each connection between nodes
- Random weight is assigned first
- Added up to get final weight for those connections

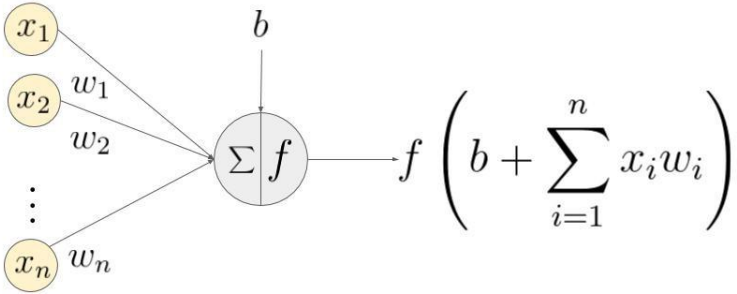


Neural Network – Bias

- Additional parameters that adjust the output with the weighted sum from input nodes
- Can affect whether or not the node can be activated

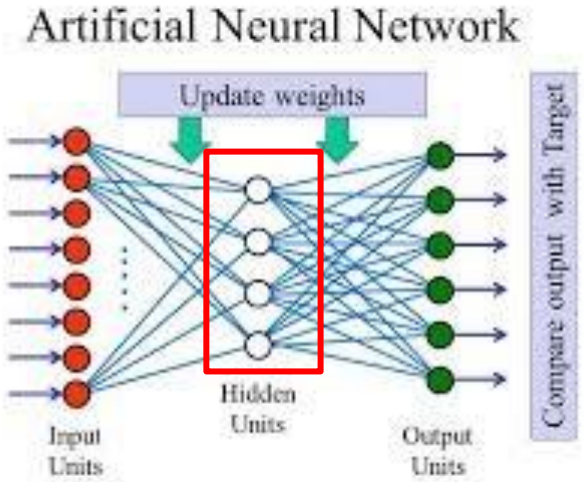


Neural Network - Activation Function



An example of a neuron showing the input (x_1 - x_n), their corresponding weights (w_1 - w_n), a bias (b) and the activation function ƒ applied to the weighted sum of the inputs.

- Algorithm that takes in the summed weights
- Used Bipolar Sigmoid
  - Range [-1,1]
  - Function is f(x) = -1 + (2/[1+e^-x])



Background

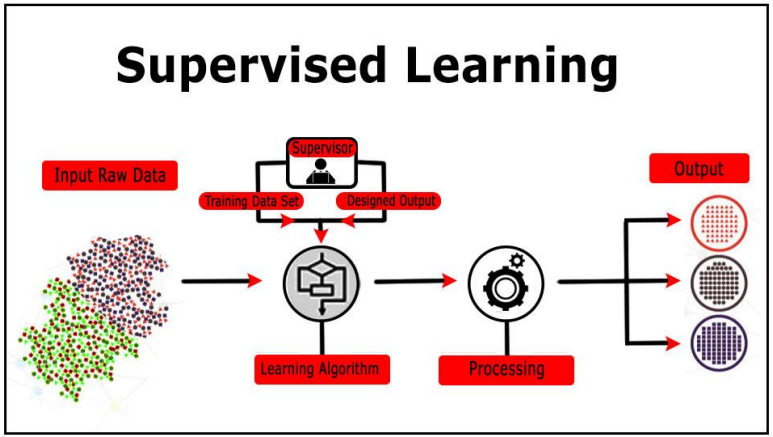
- Machine Learning
- Neural Network
- **Training**
  - **Supervised**
  - **Back-propagation**
  - **Mean Squared Error**
  - **Epoch**
- Decision Tree
- Educational Data-mining

Training

- Means giving the model a set of data to use and learn
- Set of data is training set

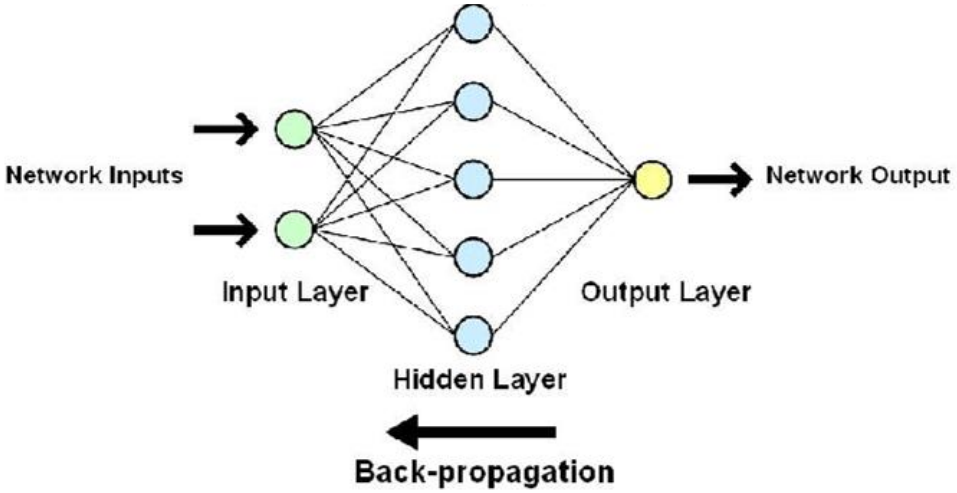
Training - Supervised

- The model is given the inputs and outputs of the training set
- Has to figure out how to adjust the weights to get correct outputs



Training - Back-propagation

- The neural network will stop before the output and send the data back to the beginning
- Allows neural network to adjust weights to have the least amount of error



Training -
Mean Squared Error

- Measure of how far neural network prediction is from the target prediction

Training - Epoch

- One full cycle through the data-set from input to output
- Relies on forward pass, backward pass, iterations, and batches
  - Forward pass goes from input to output in neural network
  - Backward pass is from output to input in neural network
  - Batch is a group of data-set that goes through forward and backward pass
  - An iteration is how many times a batch passes through the model
    - Ex: 1000 instances with batch size of 500 equals 2 iterations and 1 epoch.

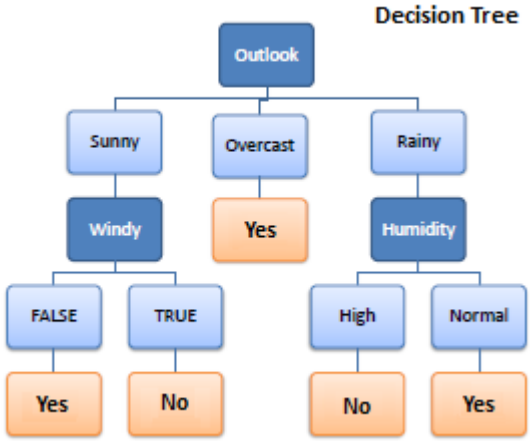
Background

- Machine Learning
- Neural Network
- Training
- **Decision Tree**
  - **Random Forest**
- Educational Data-mining

Decision Tree

- What is a tree
- Works by selecting predictors and targets
- Makes trees with picks that are the best out of the predictors

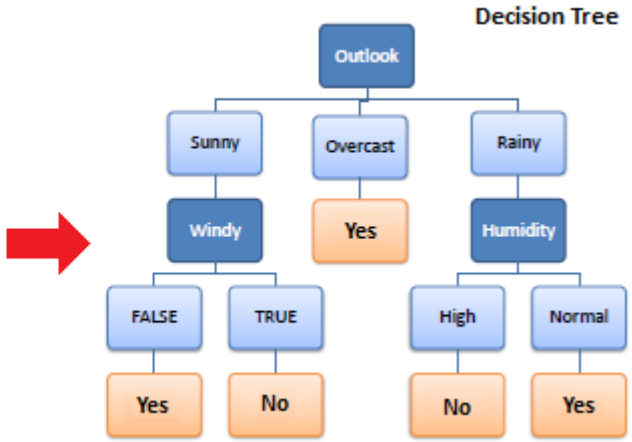
Predictors				Target
Outlook	Temp.	Humidity	Windy	Play Golf
Rainy	Hot	High	False	No
Rainy	Hot	High	True	No
Overcast	Hot	High	False	Yes
Sunny	Mild	High	False	Yes
Sunny	Cool	Normal	False	Yes
Sunny	Cool	Normal	True	No
Overcast	Cool	Normal	True	Yes
Rainy	Mild	High	False	No
Rainy	Cool	Normal	False	Yes
Sunny	Mild	Normal	False	Yes
Rainy	Mild	Normal	True	Yes
Overcast	Mild	High	True	Yes
Overcast	Hot	Normal	False	Yes
Sunny	Mild	High	True	No



Decision Tree - Random Forest

- Version of decision tree
- Take in random set of rows and column

Predictors				Target
Outlook	Temp	Humidity	Windy	Play Golf
Rainy	Hot	High	False	No
Rainy	Hot	High	True	No
Overcast	Hot	High	False	Yes
Sunny	Mild	High	False	Yes
Sunny	Cool	Normal	False	Yes
Sunny	Cool	Normal	True	No
Overcast	Cool	Normal	True	Yes
Rainy	Mild	High	False	No
Rainy	Cool	Normal	False	Yes
Sunny	Mild	Normal	False	Yes
Rainy	Mild	Normal	True	Yes
Overcast	Mild	High	True	Yes
Overcast	Hot	Normal	False	Yes
Sunny	Mild	High	True	No



Background

- Machine Learning
- Neural Network
- Training
- Decision Tree
- **Educational Data-mining**

Educational Data Mining

- Sub-group of data-mining
- Use data from educational systems to improve the educational system
- Educational data from about 900 students

Overview

- Introduction
- Background
- **Method**
- Results
- Conclusion

Method

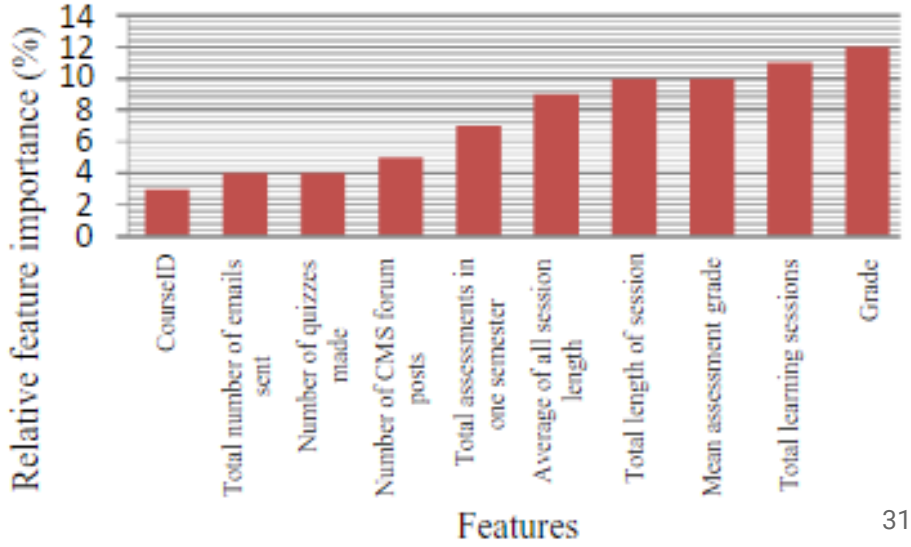
“Student Performance Prediction using Multi-Layers Artificial Neural Networks: A Case Study on Educational Data Mining” [1]

Altaf et. al

Method

- Predictors
- Architecture

Architecture	MSE	No. of Epoch	Accuracy	Error
[4x8x3]	$5.69x10^{-3}$	69	91.5	8.5
	$6.29x10^{-3}$	72	94.4	5.6
	$7.67x10^{-3}$	85	91.6	8.4
	$8.02x10^{-3}$	99	92	8
[4x12x3]	$9.49x10^{-3}$	117	96.2	3.8
	$8.99x10^{-3}$	125	96.3	3.7
	$9.02x10^{-3}$	132	97.4	2.6
	$9.79x10^{-3}$	131	97.1	2.9
[4x15x3]	$8.11x10^{-3}$	369	92.1	7.9
	$5.23x10^{-3}$	325	91.5	8.5
	$5.85x10^{-3}$	344	81.1	18.9
	$6.56x10^{-3}$	362	85.8	14.2



Overview

- Introduction
- Background
- Method
- **Results**
- Conclusion

Results

- 97.4% best accuracy in confusion matrix
- Green = correct
- Red = incorrect placement
- Gray = percent correct and incorrect in that column
- Blue = percent correct in based on the green cell

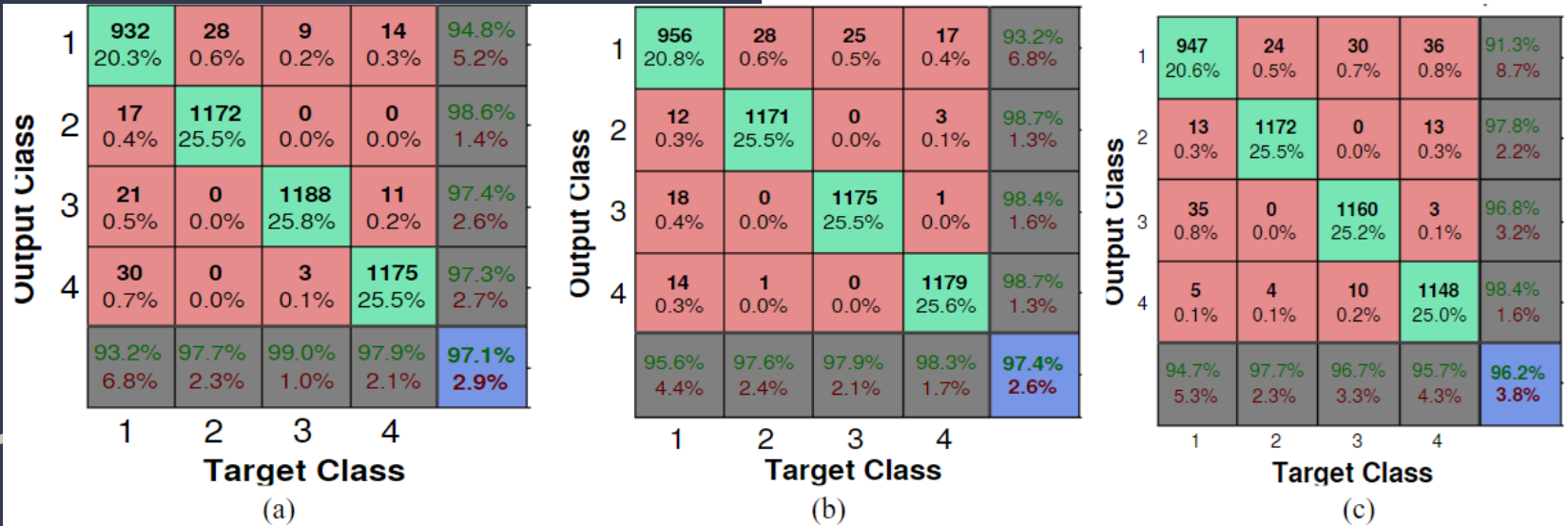


Figure 3: Confusion matrices (a) [4x8x3] (b) [4x12x3] (c) [4x15x3]

Overview

- Introduction
- Background
- Method
- Results
- **Conclusion**

Conclusion

- Best predictors is the grades
- Viable option to predict student performance

Acknowledgments

- Professor Peter Dolan
- My amazing girlfriend Hanna.

Questions?

Questions?

References

- [1] Saud Altaf, Waseem Soomro, and Mohd Izani Mohamed Rawi. 2019. Student Performance Prediction using Multi-Layers Artificial Neural Networks: A Case Study on Educational Data Mining. In Proceedings of the 2019 3rd International Conference on Information System and Data Mining (ICISDM 2019). Association for Computing Machinery, New York, NY, USA, 59–64.
- [2] Giuseppe Ciaburro and Balaji Venkateswaran. 2017. Neural networks with R: smart models using CNN, RNN, deep learning, and artificial intelligence principles, Birmingham, UK: Packt Publishing.
- [3]Toon Calders and Mykola Pechenizkiy. 2012. Introduction to the special section on educational data mining. SIGKDD Explor. Newsl. 13, 2 (May 2012), 3–6.

Image References

- “Application of Back Propagation Artificial Neural Network for Modelling Local GPS/Levelling Geoid Undulations: A Comparative Study.” https://www.semanticscholar.org/paper/Application-of-Back-Propagation-Artificial-Neural-A-G%C3%BCll%C3%BC-Yilmaz/db39fd79bb591b04d33207992f6ccde03cabd861
- “SUPERVISED LEARNING MODELS” https://www.datavedas.com/supervised-models/
- “Backpropagation” https://www.cse.unsw.edu.au/~cs9417ml/MLP2/BackPropagation.html
- “What’s the deal with Accuracy, Precision, Recall and F1?” https://towardsdatascience.com/whats-the-deal-with-accuracy-precision-recall-and-f1-f5d8b4db1021
- “Patenting Considerations for Artificial Intelligence in Biotech and Synthetic Biology – Part 2: Key Issues in Patent Subject Matter Eligibility” https://www.mintz.com/insights-center/viewpoints/2231/2020-01-patenting-considerations-artificial-intelligence-biotech

Image References

- “Use Python with Your Neural Networks” https://visualstudiomagazine.com/articles/2014/11/01/use-python-with-your-neural-networks.aspx
- “Decision Tree - Classification” https://www.saedsayad.com/decision_tree.htm
- “Build up a Neural Network with python” https://towardsdatascience.com/build-up-a-neural-network-with-python-7faea4561b31
- What Is a Confusion Matrix? https://magoosh.com/data-science/what-is-a-confusion-matrix/