

NeuroXL Classifier

Harnessing the Power of Neural Networks
for Classification of Microsoft Excel Data

A White Paper

Table of Contents

The Value of Neural Networks for Classification 1
The Advantages of NeuroXL Classifier over the Competition 3
Example: Using NeuroXL Classifier to Classify Country Data 4
Conclusion 7

The Value of Neural Networks for Classification

In finance, science, and business, analysts are often faced with the task of classifying items based on historical or measured data. Stock market analysts may wish to categorize a group of stocks as buy, sell, or hold; a cancer researcher may wish to categorize a list of tumors as benign or malignant; a mortgage analyst may wish to categorize loans as good or bad. A major difficulty faced by such analysts is that the data to be classified can often be quite complex, with numerous interrelated variables. The time and effort required to develop a model to solve accurately such classification problems can be significant.

Neural networks are a proven, widely used technology to solve such complex classification problems. Loosely modeled after the human brain, neural networks are interconnected networks of independent processors that, by changing their connections (known as *training*), learn the solution to a problem. NeuroXL Classifier software by AnalyzerXL implements *self-organizing* neural networks, which perform categorization by learning the trends and relationships within your data.

Despite their effectiveness, neural networks are often not used for classification due to their complexity and the learning required to implement them properly. NeuroXL Classifier removes these barriers by hiding the complexity of its advanced neural network-based methods while taking advantage of analysts' existing knowledge of Microsoft Excel spreadsheets. You simply supply the data, and NeuroXL Classifier implements a neural network that categorizes your data according to your preferences. Since users supply data through the familiar Excel interface, learning time is minimal, greatly reducing the interval between installing the software and performing classifications. The application is extremely intuitive and easy-to-use for beginners, not requiring any previous knowledge of neural networks, yet powerful enough for the most demanding professionals.

In summary, these are the key advantages of NeuroXL Classifier:

- Easy to learn and use
- No prior knowledge of neural networks required
- Integrates seamlessly with Microsoft Excel
- Provides proven neural network technology for highly accurate classification
- Detects relationships and trends in data that traditional methods overlook
- Lowest cost neural network classification product on the market

NeuroXL Classifier can be applied to problems in a variety of areas and industries, including:

Finance: NeuroXL Classifier's ability to handle numerous, often-interrelated variables makes it widely applicable to the financial industry. One such application is financial risk assessment, where NeuroXL Classifier can be used to categorize loan applications as good or bad. Another is stock market analysis, where a trader may wish to classify stocks as buy, hold, or sell based on historical data. Other applications include:

- Credit scoring.
- Bond ratings.
- Mortgage risk analysis.

Research Science: Researchers are often faced with the task of classifying chemicals, animals, cells, materials or other items based on measured or historical data. NeuroXL Classifier's ability to spot trends and relationships in large data sets makes it well suited for such applications. Specific examples include:

- Protein sequencing.
- Weather pattern classification.
- Air quality analysis.
- Insect gender determination.

Medicine: Neural networks have enjoyed widespread adoption in the field of medicine due to their ability to classify accurately diseases, genes, tumors, and other medical phenomena. Examples of the use of NeuroXL Classifier in this field include:

- Classification of EEG (electroencephalography) data of patients with sleep disorders.
- Tumor classification in breast cancer patients.
- Identification of genome types.

General Business: NeuroXL Classifier's advanced classification abilities and integration with Microsoft Excel make it a powerful and practical tool for solving business problems. Examples of typical applications include:

- Classification of sales prospects.
- Direct mail optimization.

The Advantages of NeuroXL Classifier over the Competition

Common complaints heard about other classification software packages are that they are too slow, too difficult to use, or that data must be re-formatted before processing. NeuroXL Classifier, using optimized neural network technology, produces classifications quickly and efficiently. It also requires no prior knowledge of neural networks, and is extremely easy-to-use. Being integrated into Microsoft Excel leverages your existing application knowledge and eliminates the need to export data and import the results. Useful classifications can be performed often just minutes after installation. Users just need to specify input and output references, perform a few mouse movements, and their classification is returned.

Neural network classification software typically costs hundreds, sometimes even thousands of dollars per license. NeuroXL Classifier is the first and only solution to offer this advanced technology for *less than fifty* dollars! The table below shows the price advantage of NeuroXL over competing products:

| Product | Price |
|------------------------------------|----------------|
| Ward Systems NeuroShell Classifier | \$445 |
| BrainMaker | \$195 |
| NeuroXL Classifier | \$39.95 |

NeuroXL Classifier also offers compatibility with Microsoft Excel-based trading software and with the entire AnalyzerXL product suite.

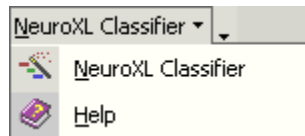
Example: Using NeuroXL Classifier to Classify Country Data

Since NeuroXL is an add-on to Microsoft Excel, it can perform classifications on your existing data already in spreadsheet form. Let's say you have a table with various economic, geographic and demographic statistics for certain countries, as shown below:

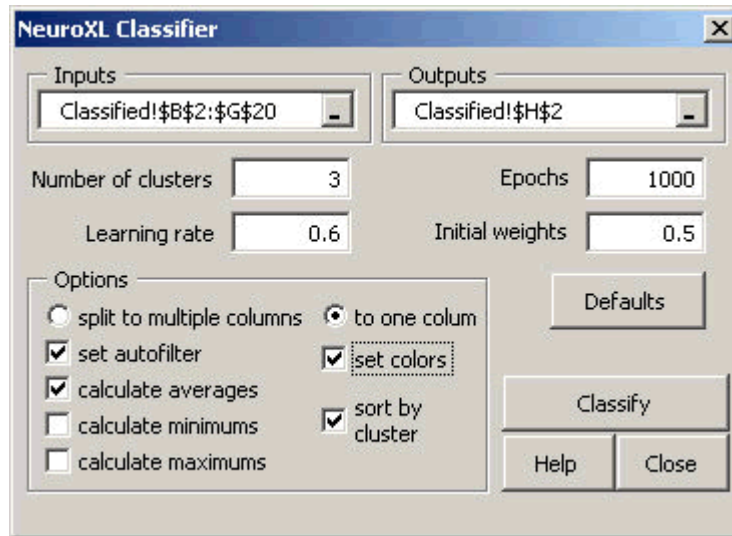
| Country | GDP Real | Population | | Area Land | Area Water | Irrigated Land |
|---------------|-------------|------------|-------------|-----------|------------|----------------|
| | Growth Rate | Population | Growth Rate | | | |
| Algeria | 3.9 | 31193917 | 1.74 | 2381740 | 0 | 5550 |
| Argentina | -3 | 36955182 | 1.16 | 2736690 | 30200 | 17000 |
| Australia | 4.3 | 19169083 | 1.02 | 7617930 | 68920 | 21070 |
| Brazil | 0.8 | 172860370 | 0.94 | 8456510 | 55455 | 28000 |
| Canada | 3.6 | 31281092 | 1.02 | 9220970 | 755170 | 7100 |
| China | 7 | 1261832482 | 0.9 | 9326410 | 270550 | 498720 |
| India | 5.5 | 1014003817 | 1.58 | 2973190 | 314400 | 480000 |
| Indonesia | 0 | 224784210 | 1.63 | 1826440 | 93000 | 45970 |
| Iran | 1 | 65619636 | 0.83 | 1636000 | 12000 | 94000 |
| Kazakhstan | 1.7 | 16733227 | -0.05 | 2669800 | 47500 | 22000 |
| Libya | 2 | 5115450 | 2.42 | 1759540 | 0 | 4700 |
| Mexico | 3.7 | 100349766 | 1.53 | 1923040 | 49510 | 61000 |
| Mongolia | 3.5 | 2650952 | 1.54 | 1565000 | 0 | 800 |
| Niger | 2 | 10075511 | 2.75 | 1266700 | 300 | 660 |
| Peru | 2.4 | 27012899 | 1.75 | 1280000 | 5220 | 12800 |
| Russia | 3.2 | 146001176 | -0.38 | 16995800 | 79400 | 40000 |
| Saudi Arabia | 1.6 | 22023506 | 3.28 | 1960582 | 0 | 4350 |
| Sudan | 3 | 35079814 | 2.84 | 2376000 | 129810 | 19460 |
| United States | 4.1 | 275562673 | 0.91 | 9158960 | 470131 | 207000 |

Now, let's say you want to divide these countries into 3 separate groups. Each group will contain countries that share similar characteristics. The steps to accomplish this are:

1. Select NeuroXL Classifier from the menu in MS Excel.



This will launch the program, bringing up the NeuroXL Classifier dialog box.



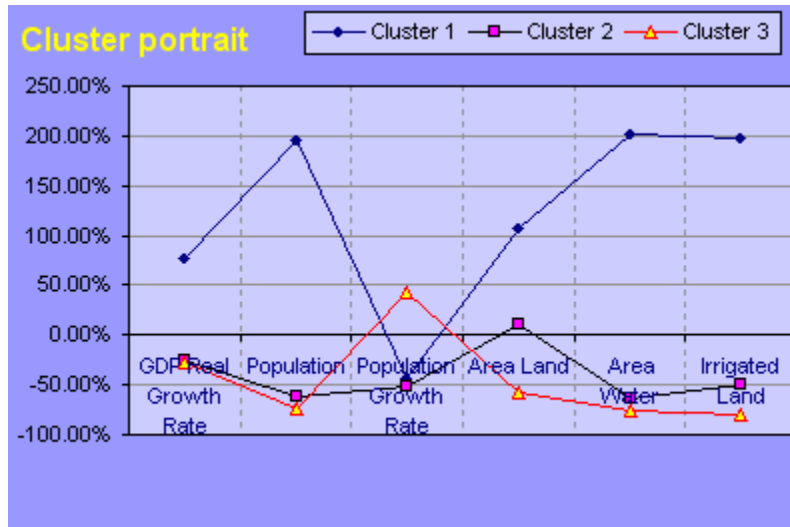
2. Specify the range of input data that you wish to be classified under “Inputs”, and then specify the cell where you would like to start outputting the classified data.
3. Set the following parameters:
 - Number of clusters: the number of clusters (categories) into which to divide the data.
 - Set autofilter: with this selected, you will be able to review the table rows being filtered.
 - Split to multiple columns: when selected, a column will be created for each cluster. For each item, if it belongs to that cluster, a “1” will be displayed in the cluster cell. Otherwise, the cell will contain a “0”.
 - To one column: when selected, a single column is created and for each item, a cluster number is inserted into the cell for that column.
 - Neural network parameters: specify the epochs, learning rate, and initial weights. In most cases, the default values need not be changed.
 - Set colors: checking this option highlights the data from each category in a different color.
 - Calculate averages: provides the averages for each category of the data items used in the categorization. In most cases, this information is extremely important. By comparing the average of a category to the total average and averages of other categories, the key characteristics of a category can be determined.
 - Calculate minimums: calculates the minimum data values for each category.

- Calculate maximums: calculates the maximum data values for each cluster
- Sort table by clusters: sorts the original table by clusters.

4. Click the “Classify” button. The results for this operation are:

| Country | GDP Real Growth Rate | Population | Population Growth Rate | Area Land | Area Water | Irrigated Land | Cluster |
|---------------|----------------------|------------|------------------------|-----------|------------|----------------------|---------------------------|
| United States | 4.1 | 275562673 | 0.91 | 9158960 | 470131 | 207000 | 1 |
| China | 7 | 1261832482 | 0.9 | 9326410 | 270550 | 498720 | 1 |
| Canada | 3.6 | 31281092 | 1.02 | 9220970 | 755170 | 7100 | 1 |
| India | 5.5 | 1014003817 | 1.58 | 2973190 | 314400 | 480000 | 1 |
| Russia | 3.2 | 146001176 | -0.38 | 16995800 | 79400 | 40000 | 1 |
| Australia | 4.3 | 19169083 | 1.02 | 7617930 | 68920 | 21070 | 2 |
| Brazil | 0.8 | 172860370 | 0.94 | 8456510 | 55455 | 28000 | 2 |
| Kazakhstan | 1.7 | 16733227 | -0.05 | 2669800 | 47500 | 22000 | 2 |
| Iran | 1 | 65619636 | 0.83 | 1636000 | 12000 | 94000 | 2 |
| Algeria | 3.9 | 31193917 | 1.74 | 2381740 | 0 | 5550 | 3 |
| Mexico | 3.7 | 100349766 | 1.53 | 1923040 | 49510 | 61000 | 3 |
| Mongolia | 3.5 | 2650952 | 1.54 | 1565000 | 0 | 800 | 3 |
| Argentina | -3 | 36955182 | 1.16 | 2736690 | 30200 | 17000 | 3 |
| Indonesia | 0 | 224784210 | 1.63 | 1826440 | 93000 | 45970 | 3 |
| Sudan | 3 | 35079814 | 2.84 | 2376000 | 129810 | 19460 | 3 |
| Saudi Arabia | 1.6 | 22023506 | 3.28 | 1960582 | 0 | 4350 | 3 |
| Libya | 2 | 5115450 | 2.42 | 1759540 | 0 | 4700 | 3 |
| Peru | 2.4 | 27012899 | 1.75 | 1280000 | 5220 | 12800 | 3 |
| Niger | 2 | 10075511 | 2.75 | 1266700 | 300 | 660 | 3 |
| | 4.68 | 545736248 | 0.806 | 9535066 | 377930.2 | 246564 | Cluster 1 average |
| | 1.95 | 68595579 | 0.685 | 5095060 | 45968.75 | 41267.5 | Cluster 2 average |
| | 1.91 | 49524120.7 | 2.064 | 1907573.2 | 30804 | 17229 | Cluster 3 average |
| | 54.80% | 82.21% | 22.67% | 57.66% | 83.12% | 80.82% | Cluster 1 average weights |
| | 22.83% | 10.33% | 19.27% | 30.81% | 10.11% | 13.53% | Cluster 2 average weights |
| | 22.37% | 7.46% | 58.06% | 11.53% | 6.77% | 5.65% | Cluster 3 average weights |
| | | | | | | Cluster 1 weight (%) | 26.32% |
| | | | | | | Cluster 2 weight (%) | 21.05% |
| | | | | | | Cluster 3 weight (%) | 52.63% |

The countries are now grouped into three distinct categories. The average data values for these categories are displayed, as well as the weight of each cluster, which represents the percentage of items belonging to the cluster. A graphical representation of each cluster can be created from the data above:



From this graph, we can determine the characteristics of each cluster. For example, cluster 1 contains countries with large populations, below average population growth, large land and water areas, plentiful irrigated lands, and above average GDP growth.

In four simple steps, we have performed a complex neural network classification of our data. NeuroXL Classifier does the work of detecting relevant patterns and trends and then grouping the data into the resulting categories. All the user needs to do is specify the inputs and set the required parameters.

Conclusion

NeuroXL Classifier is a powerful, easy-to-use and affordable solution for advanced classification of simple and complex data. By harnessing the latest advances in artificial intelligence and neural network technology, it delivers accurate and fast classifications. Designed as an add-on to Microsoft Excel, it is easy to learn and use and requires no importing or exporting of data. NeuroXL Classifier can be applied to solve problems in numerous industries and disciplines, including finance, business, medicine, and research science.

For more information on NeuroXL Classifier, please visit our Web site www.neuroxl.com or contact our support staff at support@neuroxl.com.