



## Strategic Sponsor



## Gold Sponsors



## Silver Sponsors



## Technical Partners



## Academic Partners



Wyższa Szkoła Zarządzania  
i Bankowości w Krakowie

Wyższe Szkoły Bankowe

## Media Partners



Made in Wro  
Do IT here!





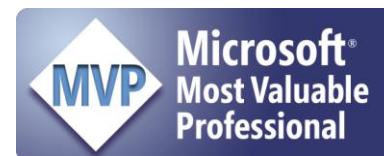
# Azure Machine Learning

## Designing Iris Multi-Class Classifier



# Marcin Szeliga

- 20 years of experience with SQL Server
  - Trainer & data platform architect
  - Books & articles writer
  - Speaker at numerous conferences
  - SQL Microsoft Most Valuable Professional since 2006
  - President of PLSSUG
  - Founder of SQLExpert
    - <http://sqlexpert.pl/>
    - <http://blog.sqlexpert.pl/>
    - [marcin@sqlexpert.pl](mailto:marcin@sqlexpert.pl)
- linkedin.com/in/marcinszeliga  
facebook.com/marcin.szeliga.18



# Session Overview

- Machine Learning overview
- Microsoft Azure overview
- Designing an Experiment using AzureML
  - Iris Multi-Class Classifier
- Deploying a Model as a service
- Monetizing Your Azure ML application



# Machine Learning Overview

- Machine learning is a discipline that emerged from the general field of artificial intelligence only quite recently
  - To build intelligent machines researchers realized that these machines should learn from and adapt to their environment
    - It is simply too costly and impractical to design intelligent systems by first gathering all the expert knowledge ourselves and then hard-wiring it into a machine
- Formal definition: “A computer program is said to learn from experience  $E$  with respect to some class of tasks  $T$  and performance measure  $P$ , if its performance at tasks in  $T$ , as measured by  $P$ , improves with experience  $E$ ” Tom M. Mitchell
- Another definition: “The goal of machine learning is to program computers to use example data or past experience to solve a given problem.”

Introduction to Machine Learning, 2nd Edition, MIT Press



# Successes and the growth of machine learning

- The first reason is rooted in its multidisciplinary character
  - Incorporated ideas from fields as diverse as statistics, probability, computer science, information theory, convex optimization, control theory, cognitive science, theoretical neuroscience, physics and more
- More important reason is the exponential growth of both available data and computer power
  - It leverages the enormous flood of data that is generated each year by satellites, sky observatories, particle accelerators, the human genome project, banks, the stock market, the army, seismic measurements, the internet, video, scanned text and so on
    - <http://www.internetlivestats.com/one-second/>



# Machine Learning Techniques

- Two primary techniques:
  - Supervised Learning
    - We are given examples of inputs and associated outputs
    - Finding the mapping between inputs and outputs using correct values to “train” a model
  - Unsupervised Learning
    - We are given inputs, but no outputs
    - Finding patterns in the input data
- Reinforcement learning (learn to select an action to maximize payoff) is difficult



# Supervised Learning

- Used when you want to predict unknown answers from answers you already have – requires data which shows the answers you can get now
- Data is divided into two parts: the data you will use to “teach” the system (data set), and the data you will use to see if the computer’s algorithms are accurate (test set)
- After you select and clean the data, you select data points that show the right relationships in the data
  - The answers are “labels”, the categories/columns/attributes are “features” and the values are...values
- Then you select an algorithm to compute the outcome
  - Often you choose more than one
- You run the program on the data set, and check to see if you got the right answer from the test set
- Once you perform the experiment, you select the best model
  - This is the final output – the model is then used against more data to get the answers you need





# Unsupervised Learning

- Used when you want to find unknown answers – mostly groupings – directly from data
- No simple way to evaluate accuracy of what you learn
- Evaluates more vectors, groups into sets or classifications
- Start with the data
- Apply algorithm
- Evaluate groups



# Machine Learning tasks

- Three common tasks
  - Classification
    - The learned attribute is categorical
  - Regression
    - The learned attribute is numeric
  - Clustering
    - Finding similar groups (clusters)



# Microsoft Azure Overview

- Setting up a Microsoft Azure Account
- Setting up an AzureML Workspace
- Accessing AzureML Studio



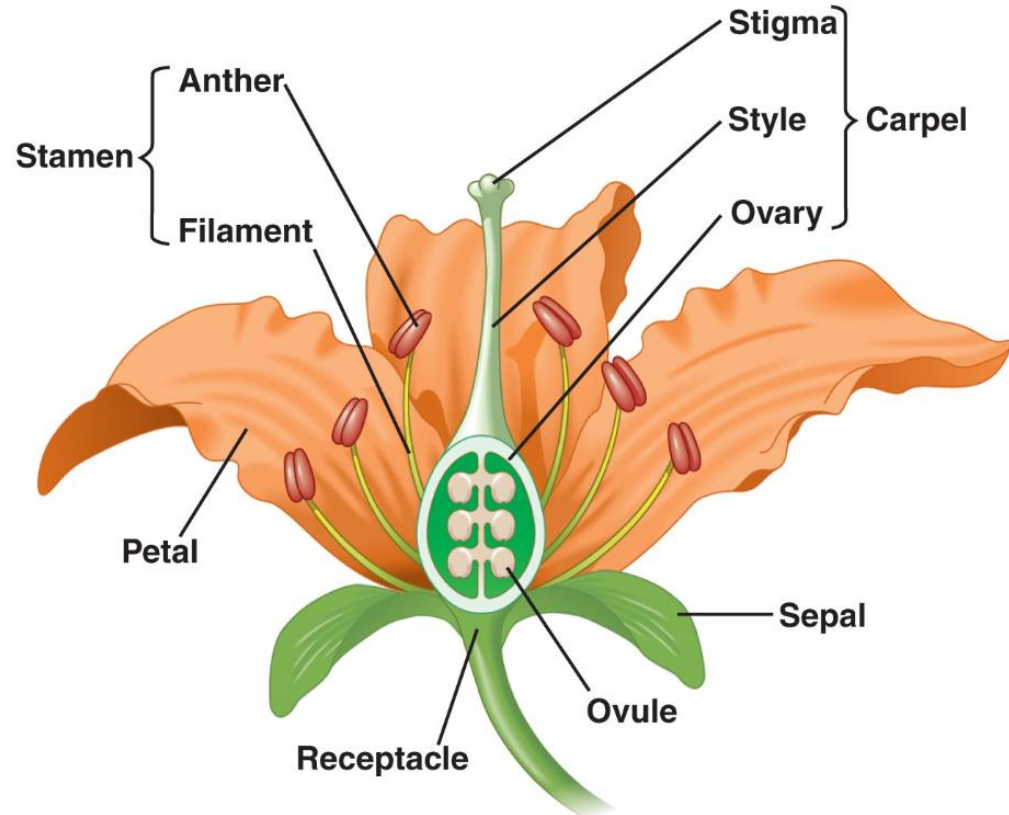
# Designing an Experiment Using AzureML

- Loading a Data Set
- Creating the Test Experiment
- Training and Scoring the Model
- Saving the Trained Model
- Creating the Scoring Experiment
- Publishing the Model
- Using the Model



# Loading a Data Set

- IRIS Dataset
  - It is perhaps the best known database to be found in the pattern recognition literature
  - The data set contains 3 classes of 50 instances each, where each class refers to a type of iris plant
    - One class is linearly separable from the other 2
    - The latter are not linearly separable from each other
  - Available from UC Irvine Machine Learning Repository
    - <http://archive.ics.uci.edu/ml/datasets/Iris>



Copyright © 2009 Pearson Education, Inc.

# Creating the Test Experiment

- Drag Iris Dataset from the «Dataset» menu item on the left and drop it in the design area
- Under the Machine Learning menu look for Initialize Model \ Classification \ Multiclass Neural Network and drop it on the design area
- Drop the Split component from the Data Transformation \ Sample and Split menu and connect the Iris Dataset to the Split Input
- Split Data between 70% for training and 30% for evaluating the model
  - Such configuration can be set in the “Properties” pane
- Drop the component from the Train \ Train Model element under Machine Learning menu
- Select the Train Model component that has been placed on the design area before and click on the «Launch columns selector» in the options area and then select the class column



# Training and Scoring the Model

- Add Score Model component from Machine Learning \ Score
- Connect Second Split output to second Score Model input
- Run the experiment
  - Model will be train used 70% of the data
  - The trained model will be used to predict the 30% of the data we already know the classification but that wasn't used in training
- Visualize the scored results, by right clicking on the Score Model output and select Visualize
- In the Visualize window, select the class column and in the «Visualization» pane, in the compare to dropdown, select «Scored Labels»



# Creating the Scoring Experiment

- Click «Create Scoring Experiment» icon
  - Saved Trained Model will replace Initialize and Train Model components
  - Web service input and output will be added
- Add Project Columns from Data Transformation \ Manipulation
  - It will be used to strip the class column from the data source and to define the correct metadata when the model will be published as a Web Service
  - Connect it with Iris Dataset and with the Score Model
  - Make sure all but class column are selected in Project Columns properties
- Run the Experiment
- Add another Project Columns connected to the Score Model
  - Strip out all the source columns and keep only the results
  - Connect it with Web service output





# Publishing the Model

- Click on the «Publish Web Service» icon
- Now the web service can be tested and give sepal and petal data as input, it will return the probability for each class and the most probable class as result
- You'll find the Web Service in the «Web Service» section of AzureML homepage
- Testing page and Excel workbook are also there
- Click Test and input new data
- See predicted values in Creating scoring experiment details



# Using the Model

- Two Web Services are available: REQUEST/RESPONSE and BATCH EXECUTION
- Both Web Services provides examples to use them with C#, R and Python
- Click API help page for REQUEST/RESPONSE Web Service
- Select R Sample Code and past it into R Studio
- Replace api\_key with key grabbed on Web Services homepage
- Input new data and run the script



# Monetizing Your Azure ML Application

- What if you needed to
  - Develop a handwriting recognition app
  - Manage a large data set
  - Use a state-of-the-art neural network
  - Deploy on thousands of devices
  - How long would that take?
- What if you could
  - Harness the power of open source
  - Combine that with enterprise-tested algorithms
  - Release that to the world
  - What could you achieve with Azure ML API Service?
- Check out the Machine Learning marketplace at [datamarket.azure.com](http://datamarket.azure.com)



# The rest is up to you





## Strategic Sponsor



## Gold Sponsors



## Silver Sponsors



## Technical Partners



## Academic Partners



Wyższa Szkoła Zarządzania  
i Bankowości w Krakowie

Wyższe Szkoły Bankowe

## Media Partners



Made in Wro  
Do IT here!

