



#ibmdevconnect

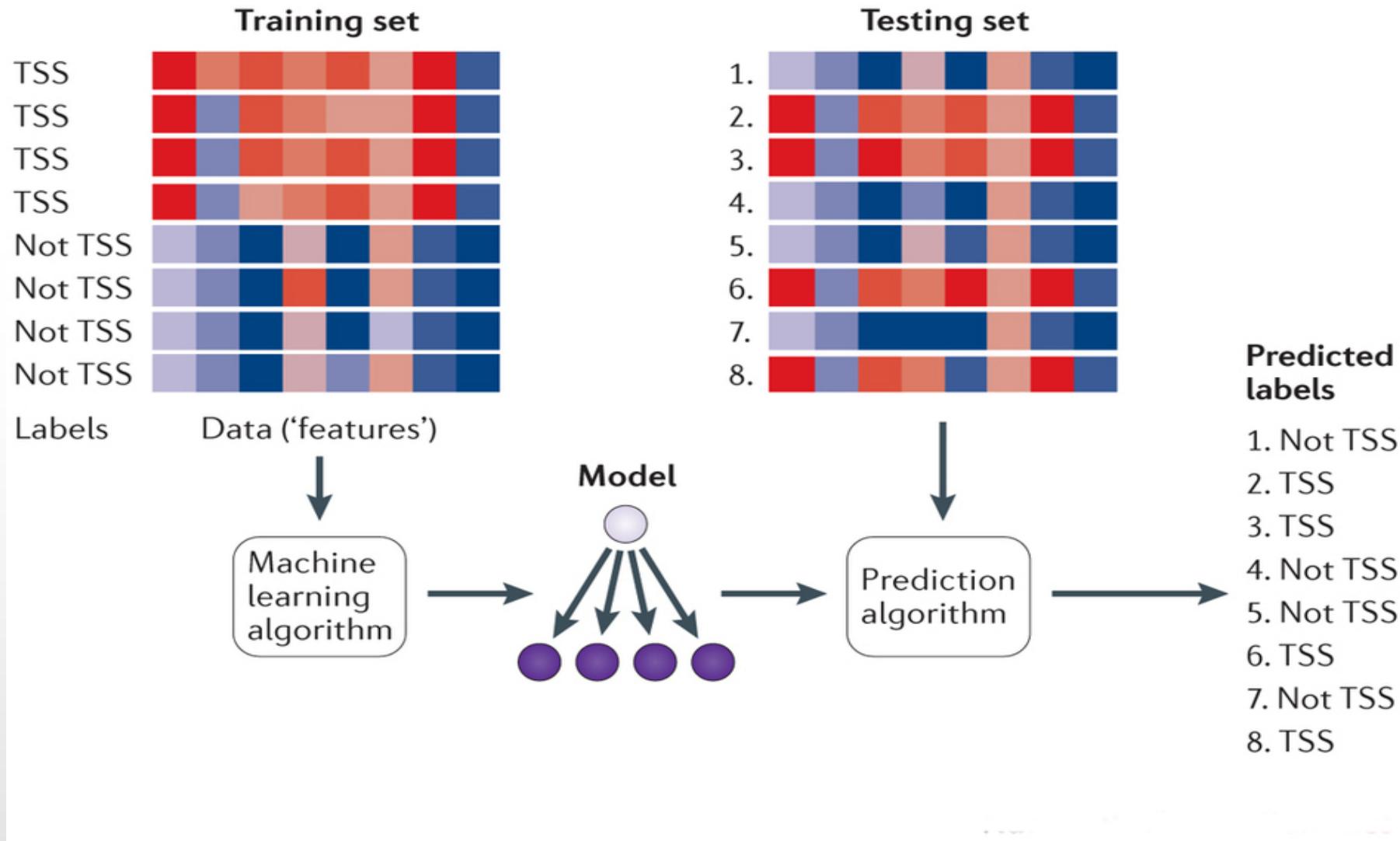
The Art of Playing with Data

Krishna Balaga
Watson and Cloud Developer
krbalaga@in.ibm.com

Takeaways

- What is Data Science and How to get Started
- Obtaining and cleaning up Datasets to Work on
- Visualizing and Analysing data using Watson Analytics tool
- Basics of Classification and Regression Modelling
- Deploying Machine Learning Models using IBM DSX Models
- Deep diving into DSX Using Notebooks

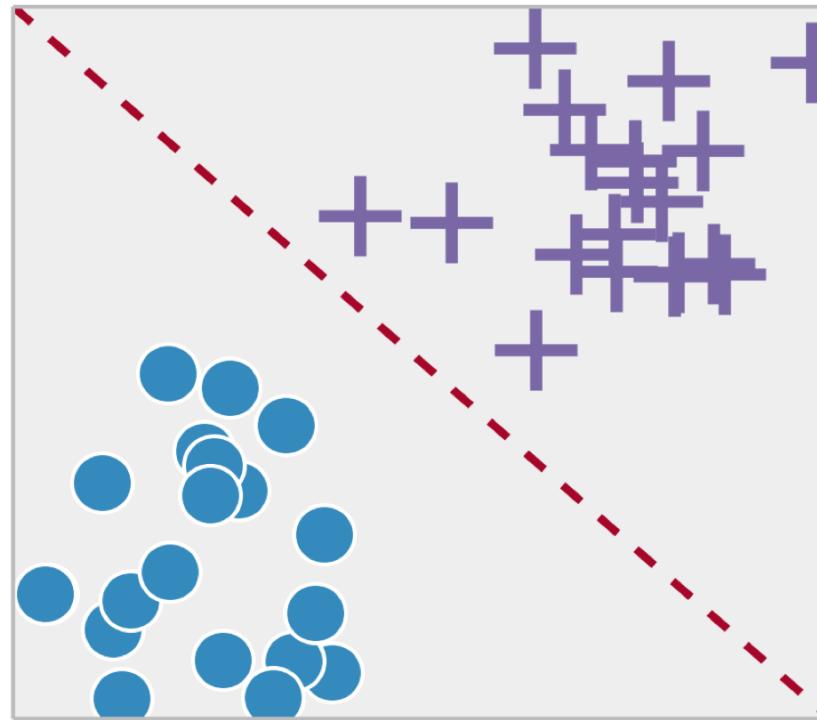
Dataset segregation



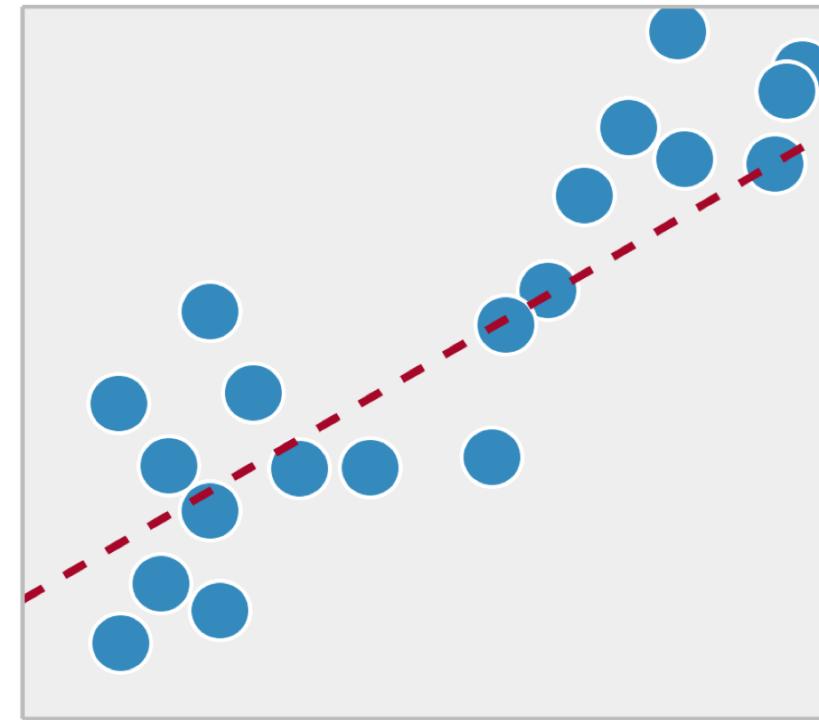
Classification: The output variable takes class labels

Regression: The output variable takes continuous values

Classification



Regression



Interesting Datasets

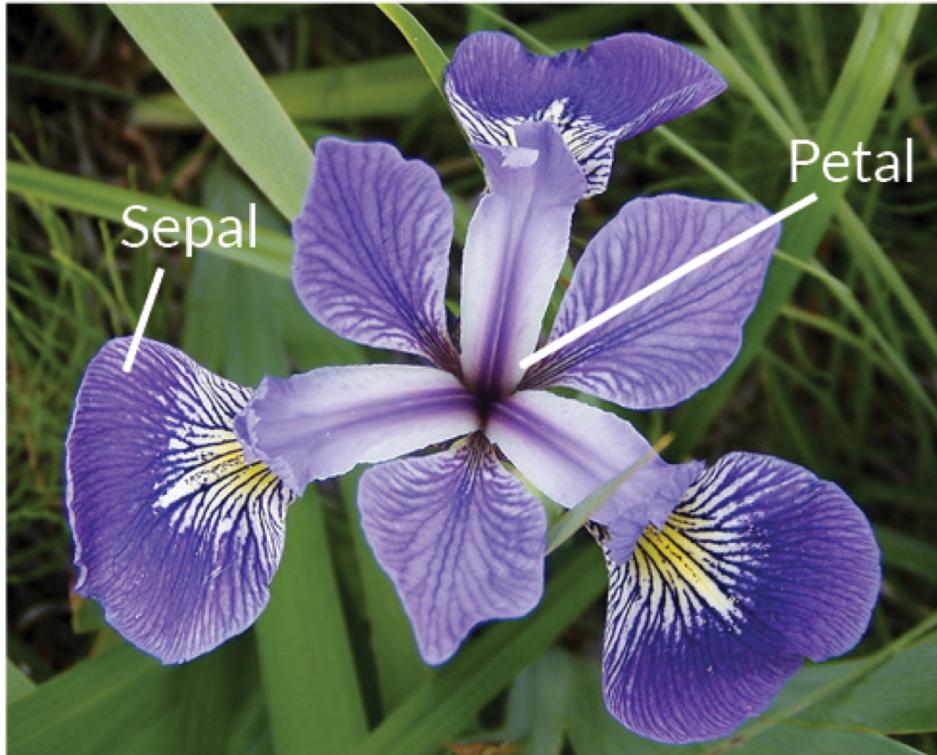
- <https://github.com/caesar0301/awesome-public-datasets>
- <https://perso.telecom-paristech.fr/eagan/class/igr204/datasets>
- <https://archive.ics.uci.edu/ml/index.php>

IRIS Dataset

- This data sets consists of 3 different types of irises' (Setosa, Versicolour and Virginica) petal and sepal length, stored in a 150x4 Table
- The rows being the samples and the columns being: Sepal Length, Sepal Width, Petal Length and Petal Width

URL : <https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>

IRIS Classifier



Iris Versicolor



Iris Setosa



Iris Virginica

Stay Connected and continue coding !



Code & instructions

<https://github.com/IBM/>

<https://github.com/IBM-Bluemix/>

[https://ibm-bluemix.github.io/#!/](https://ibm-bluemix.github.io/#/)



<https://developer.ibm.com/bluemix/>



developerWorks India - <https://developer.ibm.com/in/>

Developer Journeys - <https://developer.ibm.com/code/>



Join our Slack team and stay in touch with the experts -

<https://ibmdevconnect.slack.com>

Send in your request to - <http://ibm.biz/slackrequest>

Thank You!

Krishna Balaga

Watson and Cloud Developer
krbalaga@in.ibm.com