Visual Studio LIVE! | San Diego
EXPERT SOLUTIONS FOR .NET DEVELOPERS

An Architect's Guide to Data Science

Becky Isserman
Solution Architect
Intergen

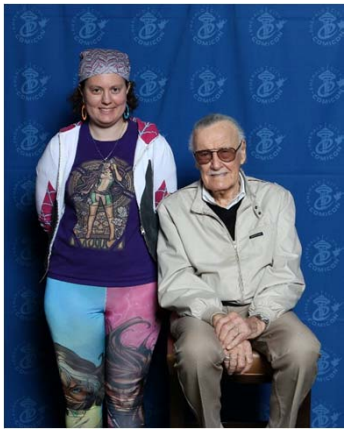
Level: Introductory

Code Again for the First Time!

Visual Studio 25 YEARS OF CODING INNOVATION

Who Am I?

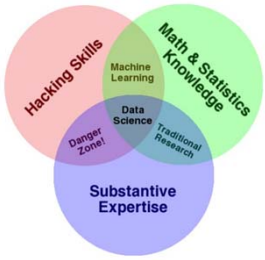
- 13 years field experience
- Former SharePoint MVP
- Azure Certified
- Former Microsoft
- Code Camp Organizer
- User Group Organizer (Eastside IoT AI ML)



Visual Studio 25 YEARS OF CODING INNOVATION

What is Data Science?

- Applies machine learning, predictive analytics, and sentiment analysis to extract critical information from collected datasets

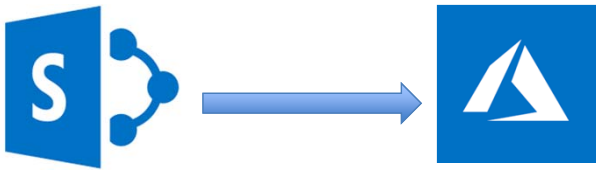


What is Machine Learning?

- The practice of using algorithms to use data to learn from it and then forecast future trends for that topic
- Predictive Analysis
- Spotting Patterns and catching hidden insights based on perceived data

From SharePoint to Machine Learning

- How do I get from Point A to Point B?
- Emotional IQ
- 2-4 years of Azure Architecture + Certs
  - 70-533
  - 70-534
  - 70-532
- PowerBI + Azure SQL



Data Analytics

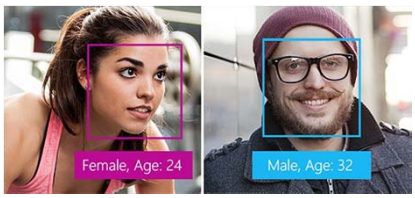
- Processing and examining data sets in order to draw conclusions about information contained in a system
- Power BI / Excel
- Charts and graphs may uncover waste
- Credit Card data may show how much money wasted monthly in Excel

Cognitive Services APIs

- Pre-made algorithms with REST endpoints that can be built into applications
- 5 areas (Vision, Speech, Language, Knowledge, Search)
- Some examples:
  - Face API
  - LUIS (Language Understanding)

Face API Demo

- https://azure.microsoft.com/en-us/services/cognitive-services/face/#detection
- https://azure.microsoft.com/en-us/services/cognitive-services/face/#recognition



Where do I start?

- Understand Cloud Systems
- Get to know your data
- Know how to process and clean the data
- Look at pre-made systems like Cognitive Services
- Learn some statistics and linear algebra

Why use machine learning?

- Identify if you need machine learning development
- Choose the most suitable type of machine learning
  - Supervised
  - Unsupervised
  - Reinforcement
  - Deep Learning

- Don’t underestimate data pre-processing and cleaning
- Choose between pre-made APIs and rolling your own

Free Courses

- AI School - https://aischool.microsoft.com/
- Coursera – http://www.coursera.org
- Khan Academy - https://www.khanacademy.org/math/linear-algebra

Paid Courses

- Udemy Top Picks – sometimes $10 a course
  - Python for Data Science and Machine Learning Bootcamp by Jose Portilla
  - Learning A-Z Courses by Kirill Eremenko
  - Data Science, Deep Learning, an Machine Learning with Python Hands on Learning by Frank Kane
- Coursera Certificates/Paid Courses - $49 a month

Certifications

- Microsoft Professional Program in AI ($990) - https://www.edx.org/microsoft-professional-program-artificial-intelligence
- Andrew NG Coursera Course ($49 a month)
- Master's Degree

Tools

- Jupyter Notebooks
- Scikit-learn
- Google Cloud ML Engine
- Azure Machine Learning Services/ML Studio
- Amazon Sagemaker

Data Bricks

- Apache Spark based cloud system
- Can house large amounts of data
- MLFlow allows for the building, testing, and deployment of models
- Integrated with Azure and PowerBI
- Also included in AWS
- Easily use Python and R

Data Bricks

- https://go.databricks.com/hubfs/notebooks/Pop._vs._Price_LR.html

Questions

- ???

References

- Data Science vs Data Analytics vs Data Engineer: https://www.simplilearn.com/data-science-vs-data-analytics-vs-machine-learning-article
- Jupyter Notebooks: http://jupyter.org/
- Scikit-learn: http://scikit-learn.org/

Contact Info

- Twitter: @undiscovereddev
- Personal Email: rebecca.Isserman@gmail.com
- Work Email: rebecca.Isserman@teamintergen.com
- Company Site: http://www.intergen.co.nz