

Introduction to Lifecycle of Data Science & IBM Model Asset Exchange

IBM Developer

Saishruthi Swaminathan | Developer Advocate | IBM San Francisco
Juanita Dion | Sr. Program Strategy Lead | IBM San Francisco

What can you take away from this workshop?

- Know the process on how to make data **speak right**

- How IBM Model Asset Exchange (MAX) can help building a desired data model in a timely and efficient manner?



Photo by Flickr user [CyperHades](#), used under Creative Commons Licensing.



Agenda

- 
- 1 • What is Data Science? Why is it important?
 - 2 • The Data Science Process | Pipeline
 - 3 • Challenges in Data Model Development
 - 4 • Introducing IBM Model Asset Exchange
 - 5 • Demo on IBM Model Asset Exchange
 - 6 • Summary
 - 7 • Workshop



What is Data Science?



Breaking down data science definition

Goal

- Find solution to the business problem

How?

- Transforming problems to well-posed questions

Using?

- Mathematics, programming and scientific method

Finally?

- Communicate results and its business impact



Business Problem : Employee Attrition

Is it a true statement?

Is there other hidden reasons ?

And|Or,

[Press Center](#) / [Press Releases](#) / 2018-01-10

GLASSDOOR SURVEY FINDS MORE EMPLOYEES EXPECTED TO QUIT IN UPCOMING YEAR, WITH SALARY CITED AS TOP REASON

New Data Shows Employee Salary Expectation Key Motivation for Resignation; One-Third of Employers Anticipate More Employees Leaving in 2018

MILL VALLEY, CALIF. (January 10, 2018) - Glassdoor®, one of the world's largest job and recruiting sites, today released new data¹ which reveals that 35 percent of hiring decision makers expect more employees to quit over the next 12 months. The survey, conducted among 750 hiring decision makers (those in recruitment, HR and responsible for hiring) in the U.S. and UK, also finds that nearly half (45 percent) note that salary is the top reason for employees changing jobs, followed by career advancement opportunities, benefits and location.



Business problem to Data science question

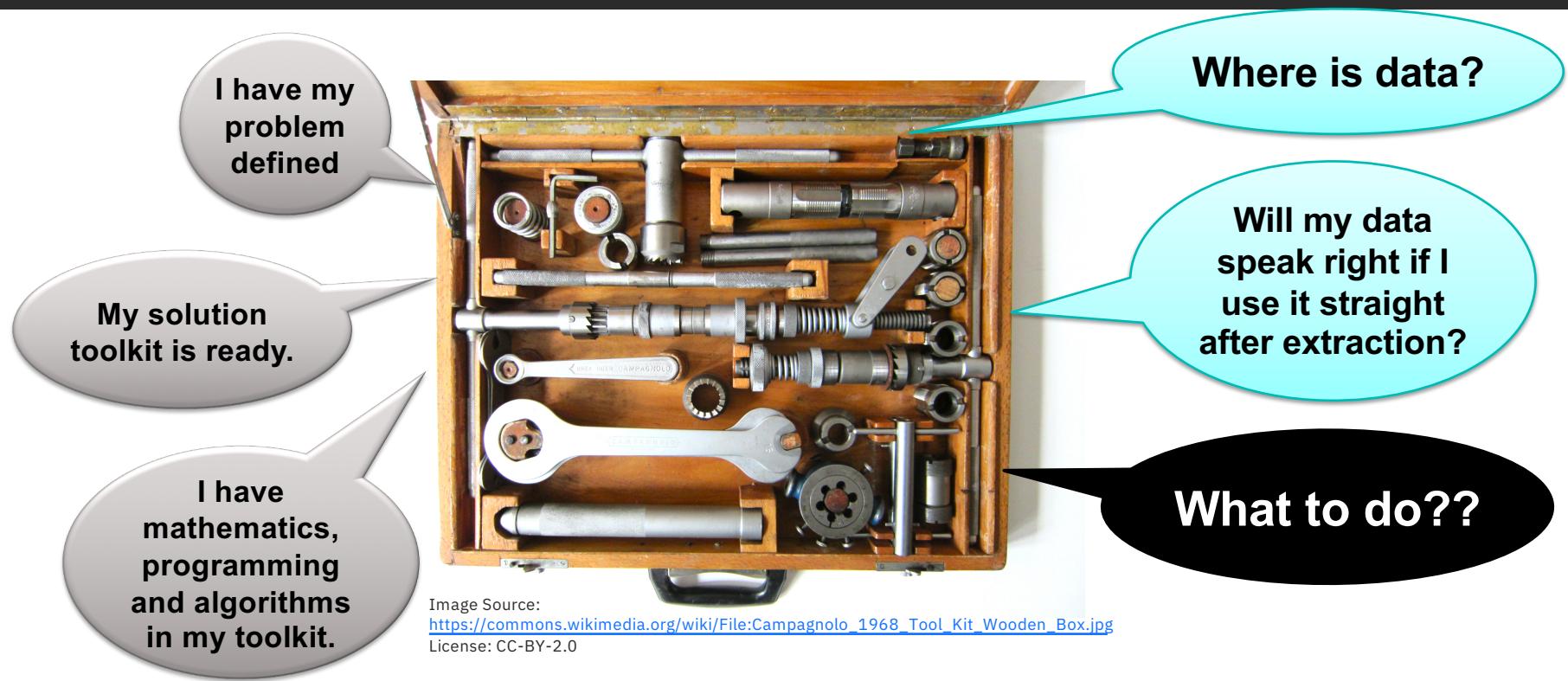
Can we find the factors having significant contribution to Employee Attrition?

Can we predict if an employee is likely to quit or not?

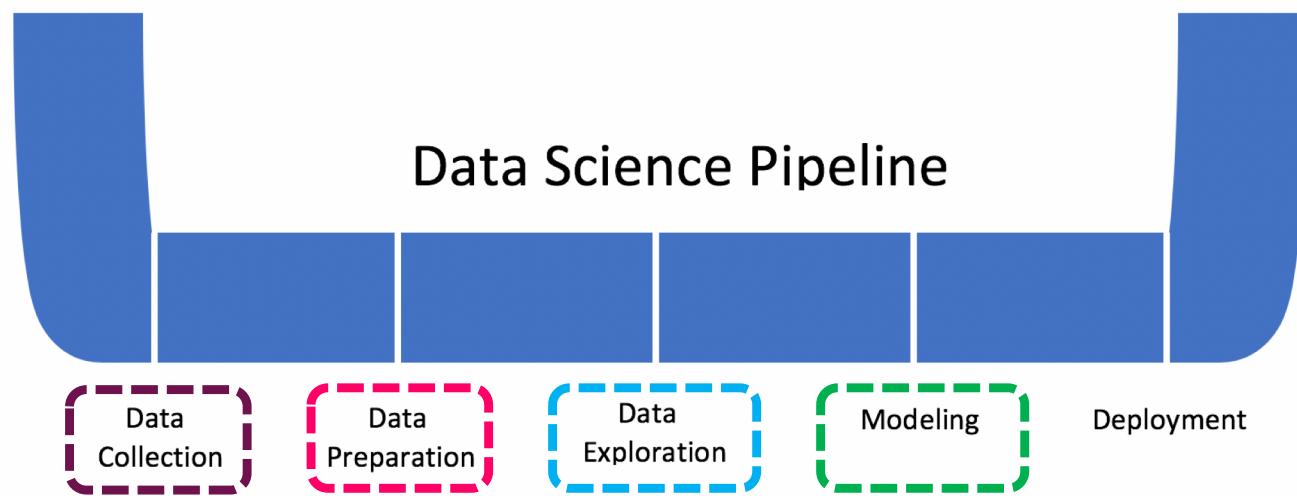
This will help HR to intervene on time and remedy the situation.



What next?



Data Science Pipeline



Challenges in data model development:

- Know what are the available models for the problem
- Pick out the best working model
- Solve installation and set-up issues
- Get the code up and running
(not all the models run in one go)



IBM Model Asset Exchange

Model Asset Exchange:

- One place for all **open source** deep learning models.
- Wide variety of domains

Provides:

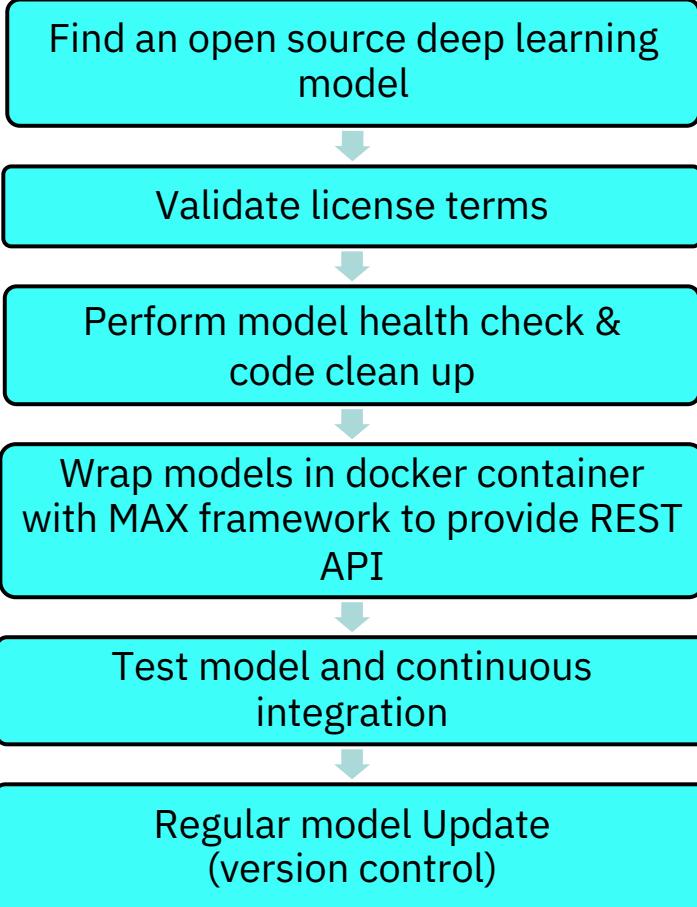
- Tested **code** and **IP**
- Both **trainable** and **trained** version

Started with the goal of:

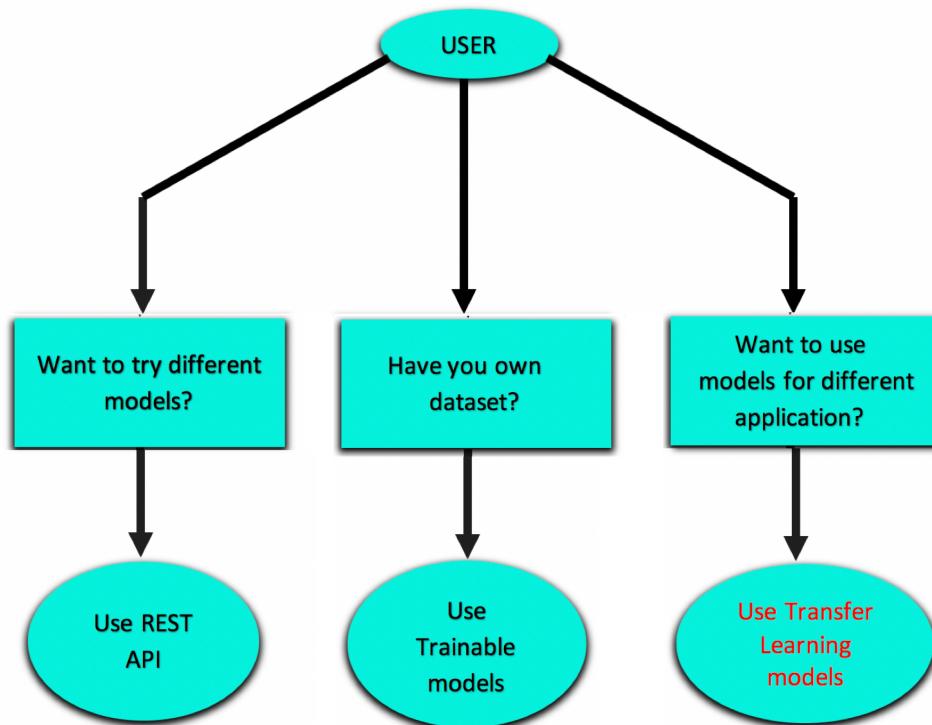
- Making **deep learning models** easy to consume.



Model Asset Exchange Work Flow



The Next Steps...

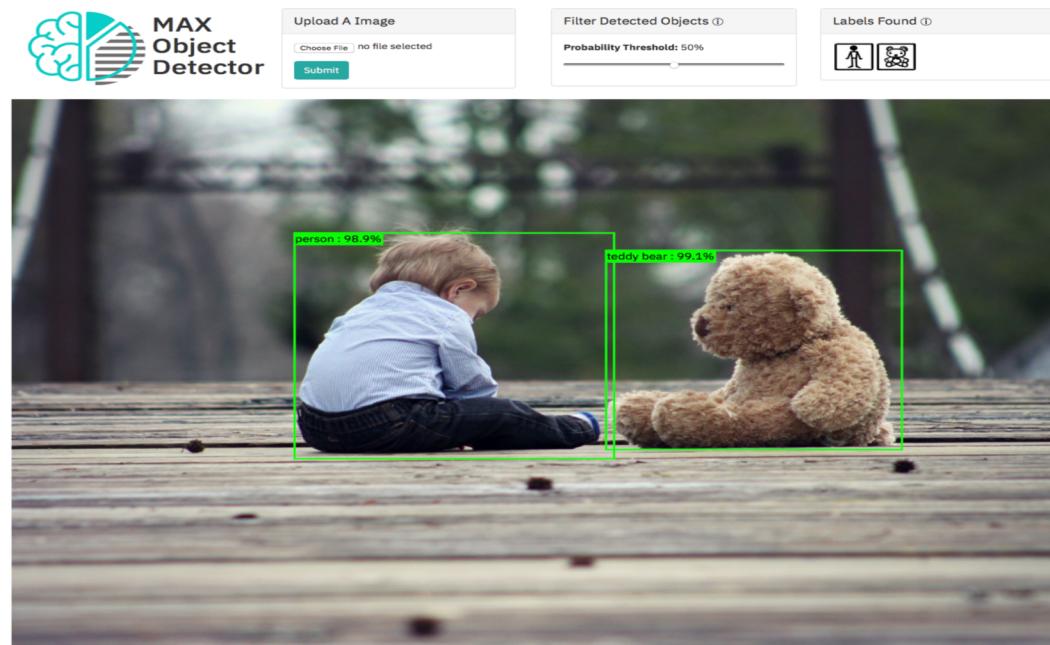


Demo of IBM Model Asset Exchange

Model : Object Detector

Model Asset Exchange link : [MAX](#)

Link to the model : [MAX-Object-detector](#)

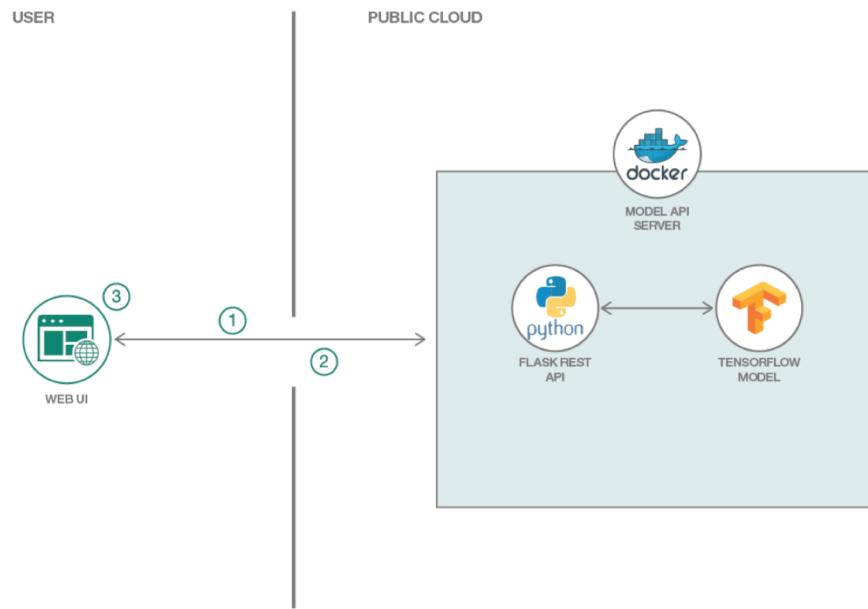


Summary

- 1 • Data science helps in resolving the business problem
- 2 • To get answers for data science question, use data pipeline
- 3 • Challenges in model development and how MAX comes to rescue
- 4 • Free, open-source models categorized into variety of domains
- 5 • Tested code and IP
- 6 • Build and deploy web service in less than 30 seconds
- 7 • Start training on Watson Studio in minutes



Workshop – Create a web app using MAX Object Detector



- User uses Web UI to send an image to Model API.
- Model API returns object data and Web UI displays detected objects.



To-do 1: Prepare input data for prediction

```
# TODO R1: review inference request payload
# Required inference request parameter: image (JPG/PNG encoded)
files = {
    'image': image_encoded.tostring(),
    'Content-Type': 'multipart/form-data',
}

# Optional inference parameter: threshold (default: 0.7, range [0,1])
data = {'threshold': '0.5'}
```



To-do 2: Send input image to Model API

```
# TODO T1: replace model URL placeholder  
model_url = args.ml_endpoint.rstrip('/') + '**TODO**'  
  
# Send image file form to model endpoint for prediction
```

Fill **TODO** with model URL which is '**/model/predict**'



To-do 3: Retrieve results from Model API

```
# TODO T2: uncomment next line and replace placeholder
# result = output_data['**TODO**']

if len(result) == 0:
    msg = 'No objects detected, try uploading a new image'
    return render_template('index.html', error_msg=msg)
```

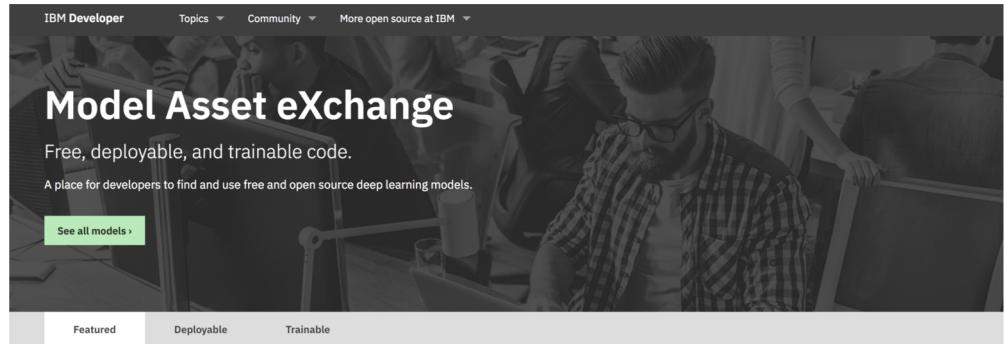
Model API returns predictions in JSON format. To retrieve result, replace **TODO** With JSON key **'Predictions'**.



Resources



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

A screenshot of the IBM Model Asset eXchange website. The header includes "IBM Developer", "Topics", "Community", and "More open source at IBM". The main title is "Model Asset eXchange" with the subtitle "Free, deployable, and trainable code. A place for developers to find and use free and open source deep learning models." Below the title are buttons for "See all models", "Featured", "Deployable", and "Trainable". The background features a grayscale photo of a developer working at a computer.

<https://developer.ibm.com/exchanges/models/>



<https://github.com/CODAIT/max-workshop-app-python>



Women in Tech Series | November 13, 2018 / © 2018 IBM Corporation



saishruthis



SSaishruthi



saishruthi.tn@ibm.com

juanita_dion

dionj

dionj@us.ibm.com



Women in Tech Series | November 13, 2018 / © 2018 IBM Corporation

