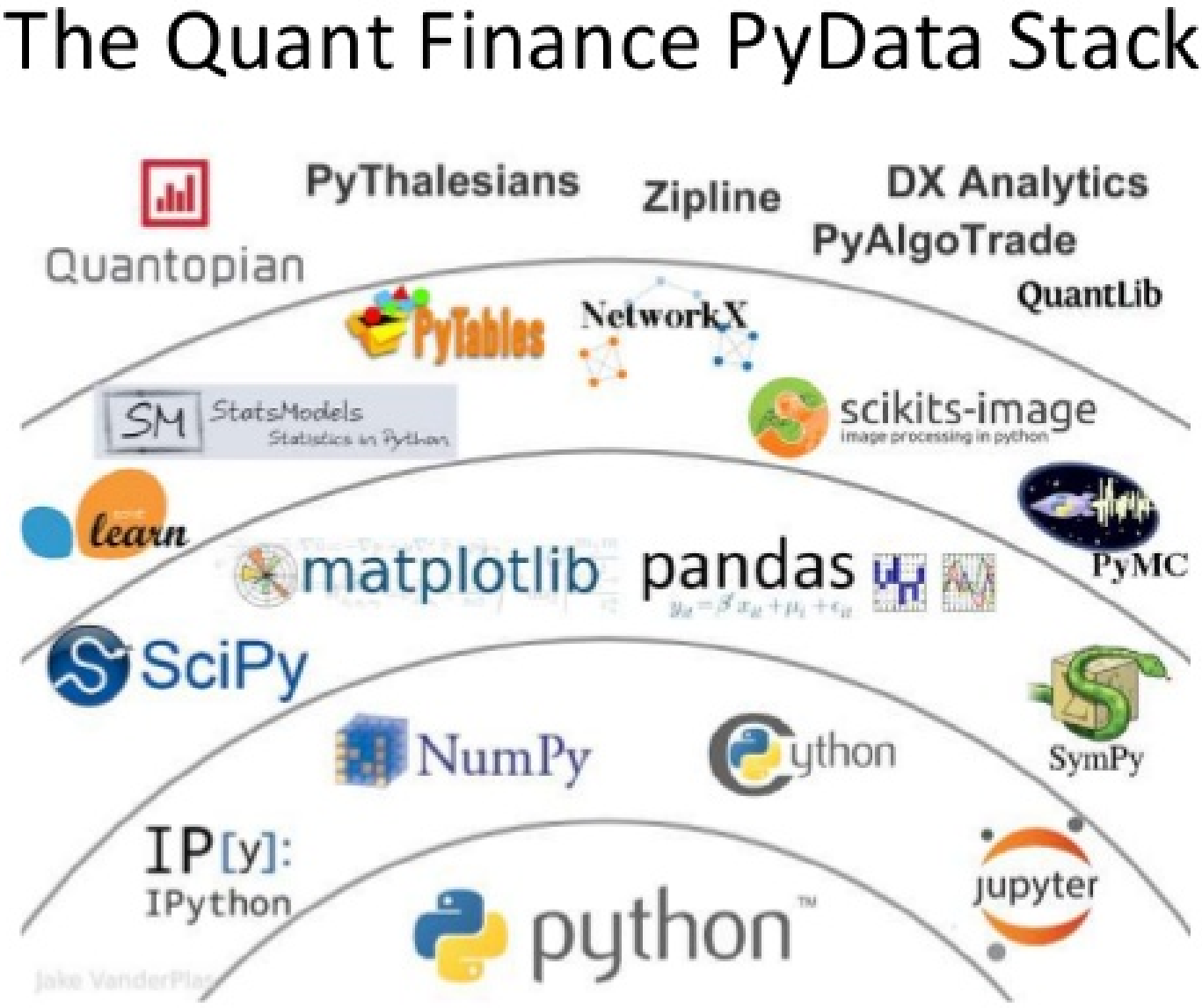
Data Analysis e Machine Learning in Python

Pietro Battiston

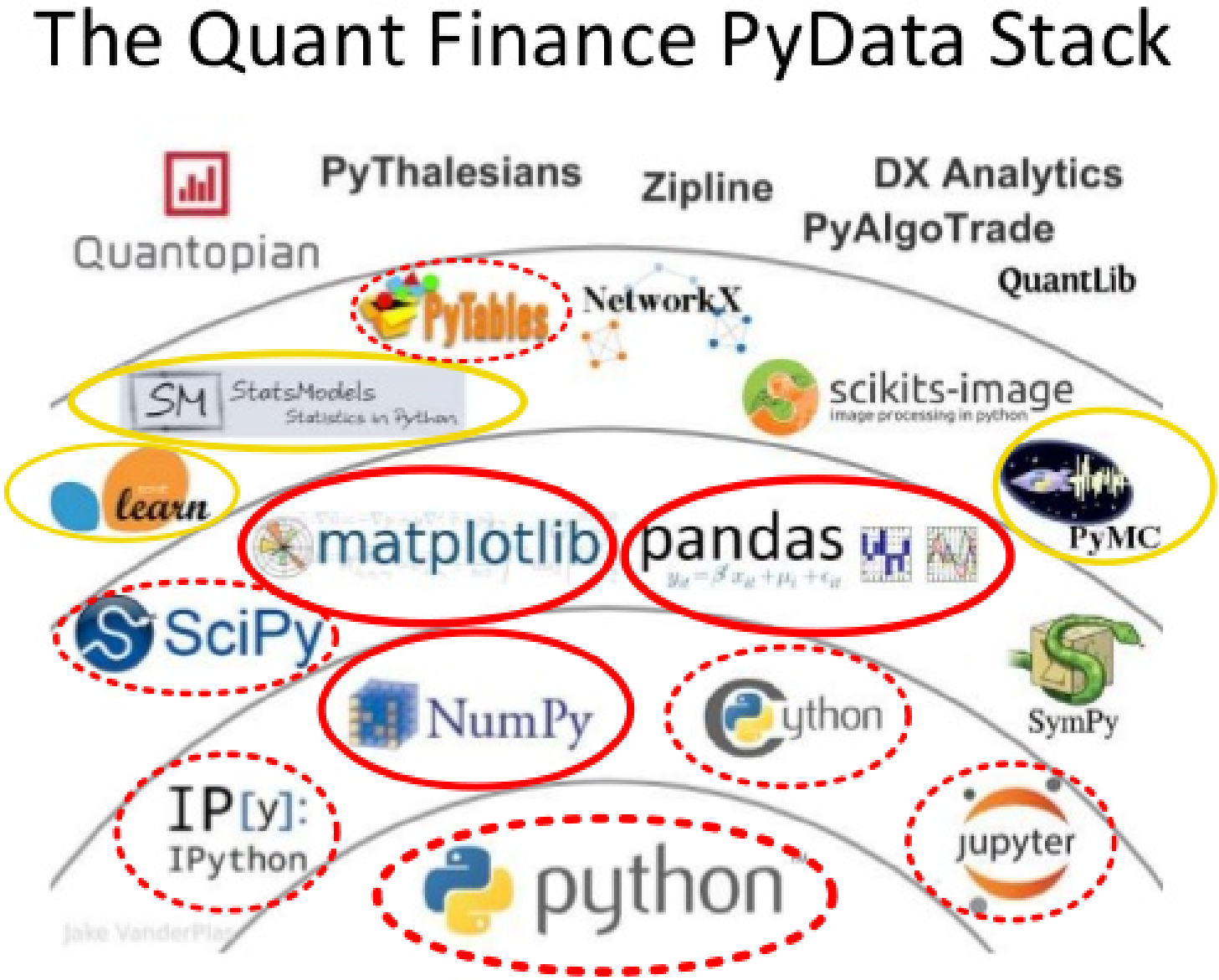
Pisa, 29 ottobre 2017
Linux Day

Orientiamoci



- Source: [Jake VanderPlas: State of the Tools]
  - (https://www.youtube.com/watch?v=5GIINDD7qbP4)

Orientiamoci

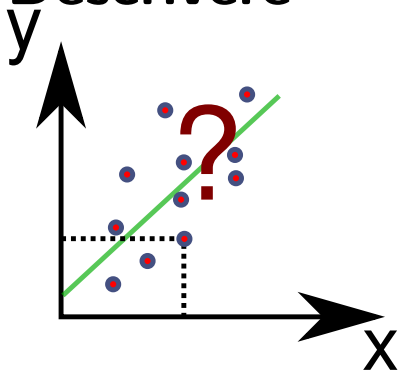


- Source: [Jake VanderPlas: State of the Tools]
  - (https://www.youtube.com/watch?v=5GIINDD7qbP4)

Cosa possiamo fare con dei dati?

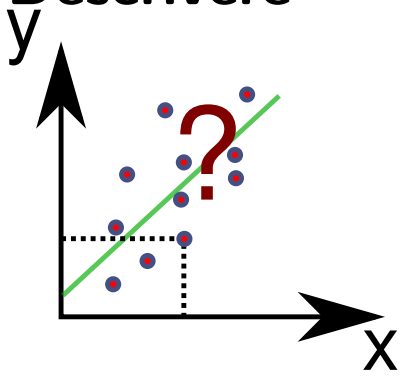
Cosa possiamo fare con dei dati?

- **Descrivere**

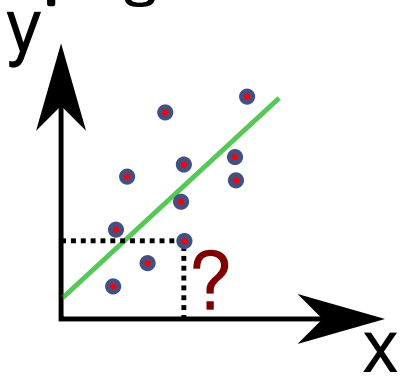


Cosa possiamo fare con dei dati?

- **Descrivere**

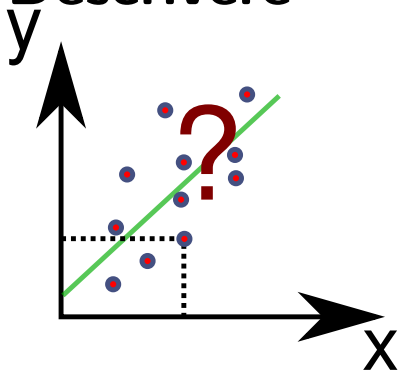


- **Spiegare**

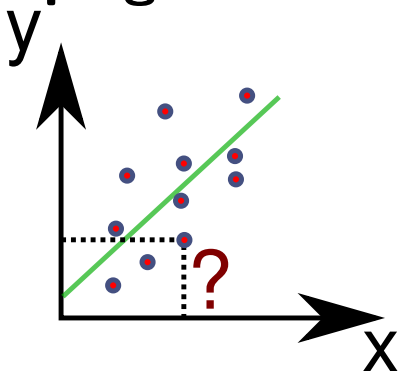


Cosa possiamo fare con dei dati?

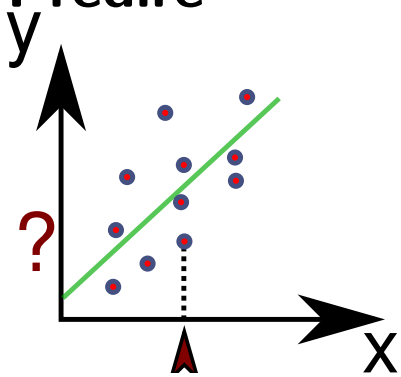
- **Descrivere**



- **Spiegare**



- **Predire**



Intanto: i dati

pandas - libreria per la manipolazione dati "etichettati"

- Source: [Jake VanderPlas: State of the Tools]
  - (https://www.youtube.com/watch?v=5GINDD7qbP4)

Intanto: i dati

pandas - libreria per la manipolazione dati "etichettati"

~~Copiata da~~ ispirata a **R**

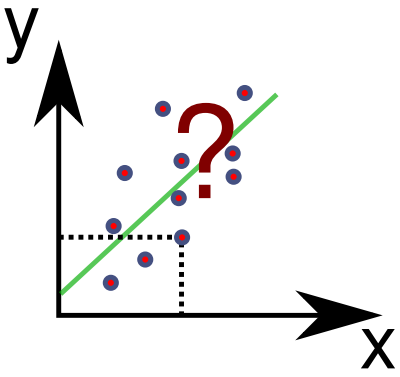
Intanto: i dati

pandas - libreria per la manipolazione dati "etichettati"

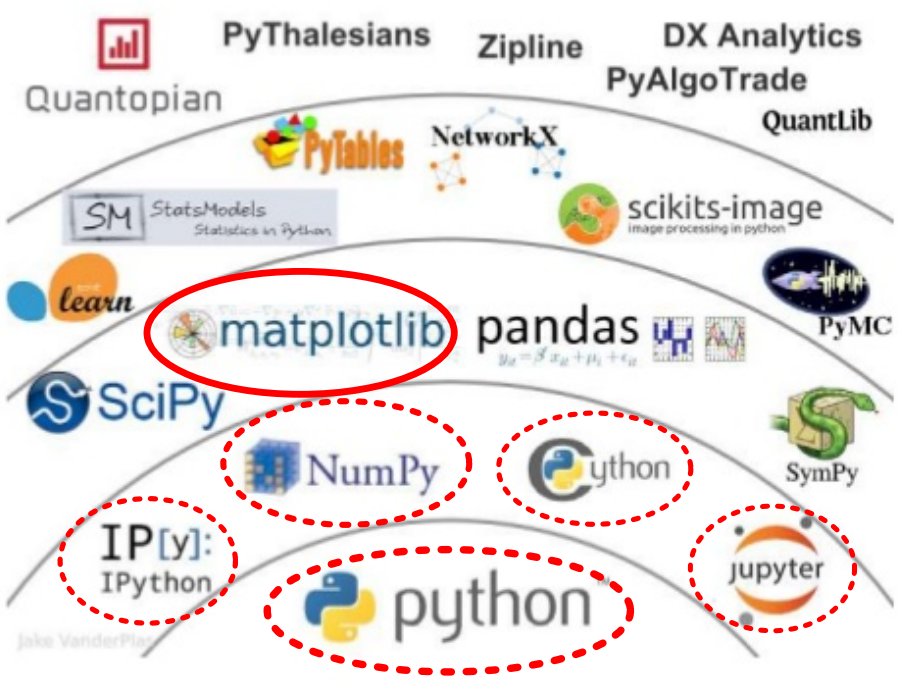
~~Copiata da~~ ispirata a **R**

Codice!

Descrivere

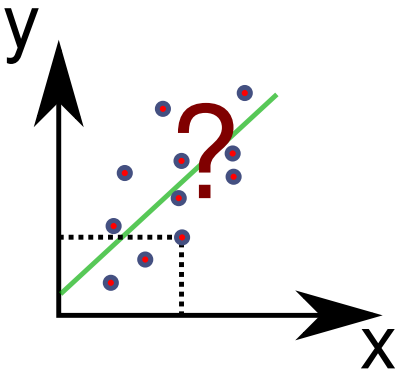


matplotlib - libreria per la *visualizzazione* dei dati

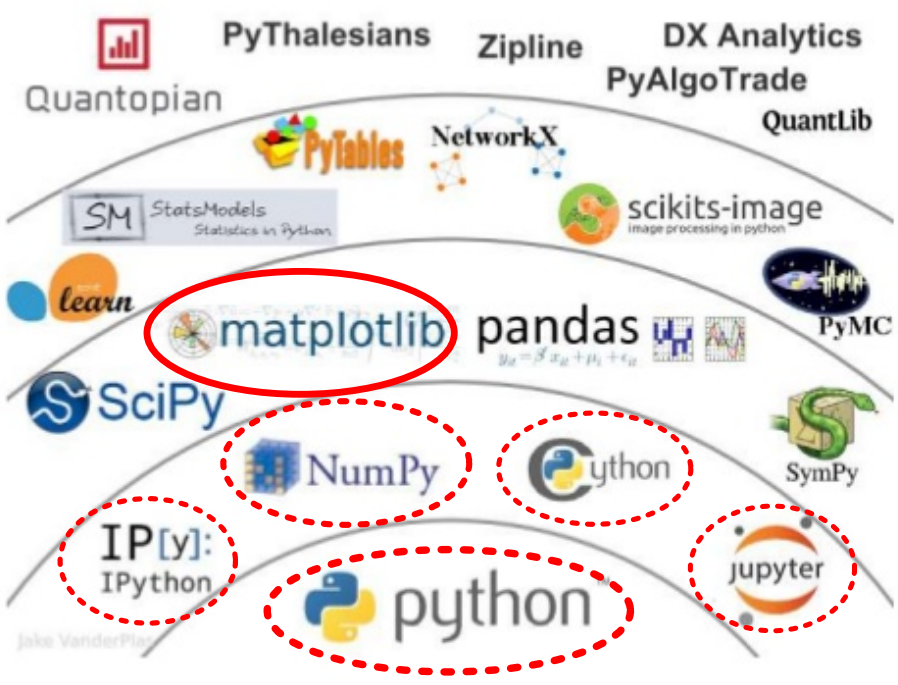


- Source: [Jake VanderPlas: State of the Tools] (https://www.youtube.com/watch?v=5GIlNDD7qbP4)

Descrivere

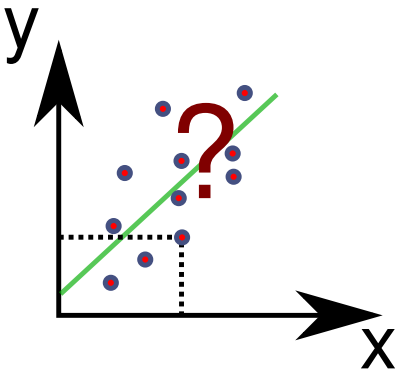


matplotlib - libreria per la *visualizzazione* dei dati

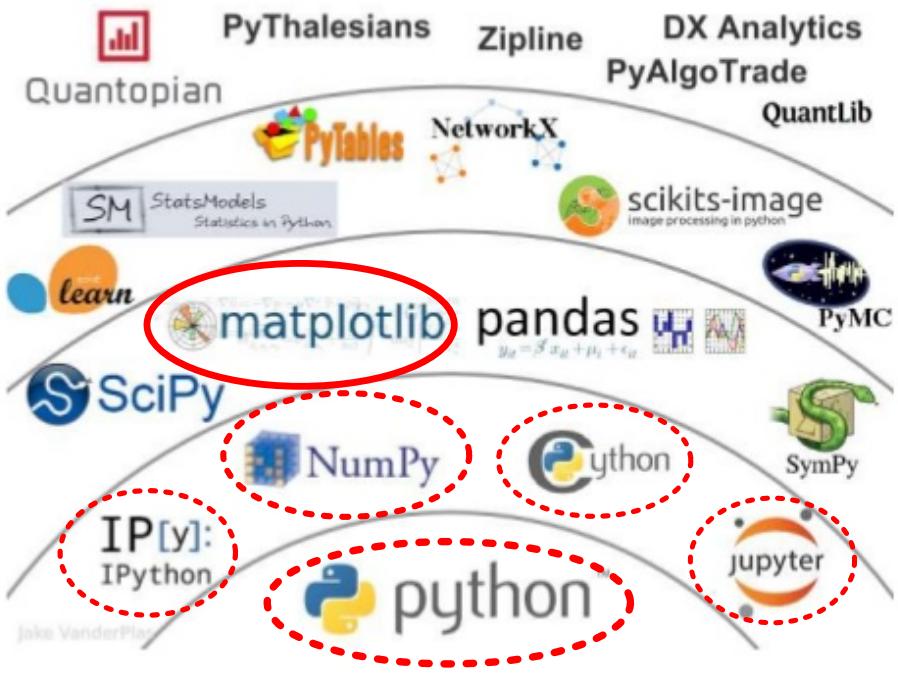


bokeh - sul web, **seaborn** - più alto livello

Descrivere

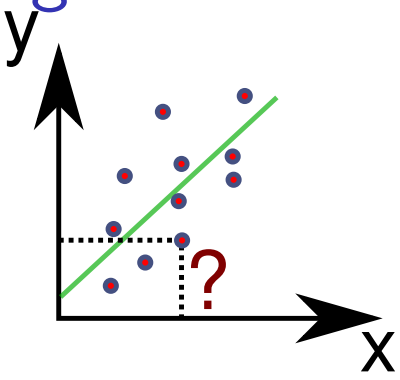


matplotlib - libreria per la *visualizzazione* dei dati

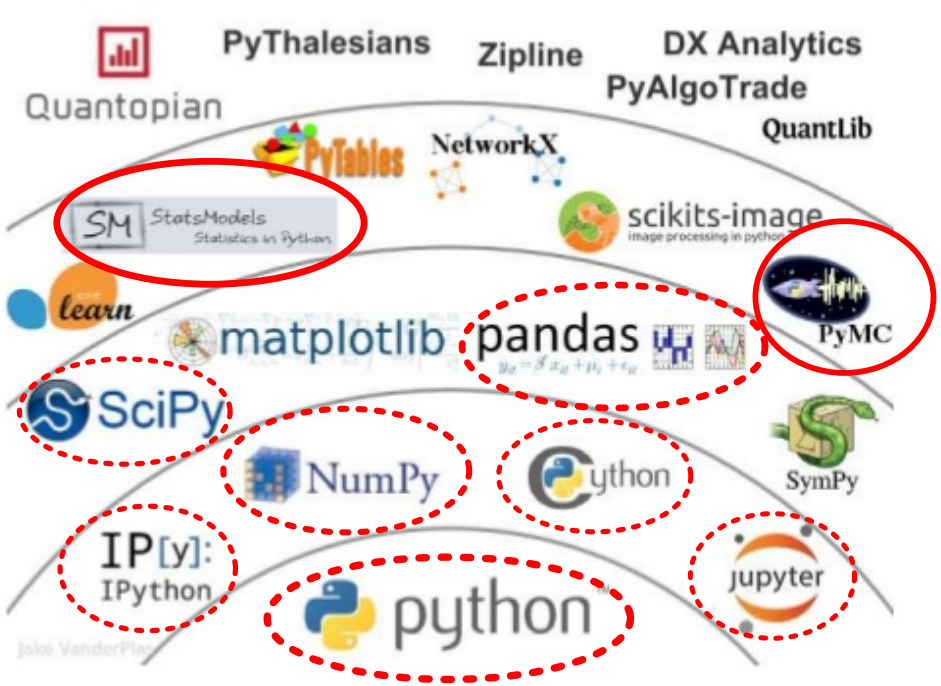


bokeh - sul web, **seaborn** - più alto livello **Codice!**

Spiegare

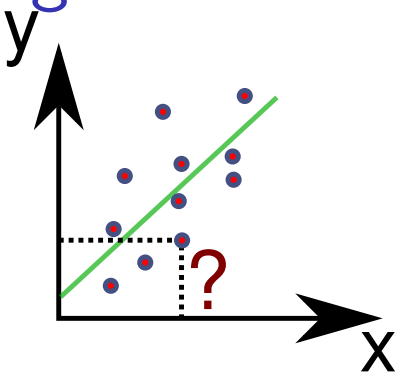


statsmodels - libreria per l'analisi statistica *classica*

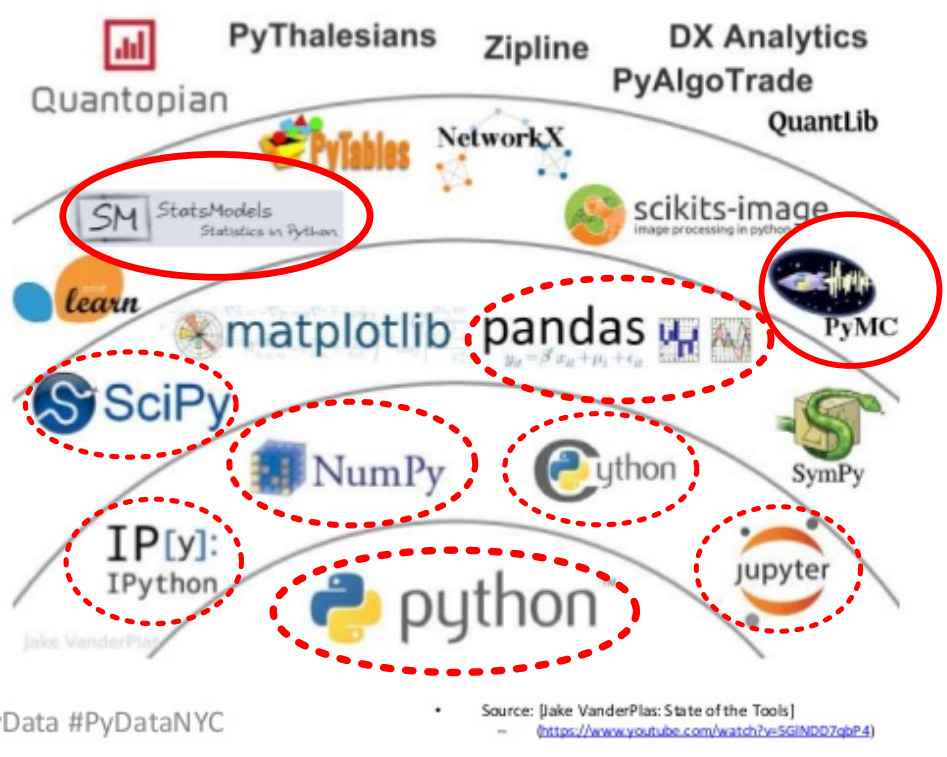


- Source: [Jake VanderPlas: State of the Tools] (https://www.youtube.com/watch?v=5GINDD7qbP4)

Spiegare

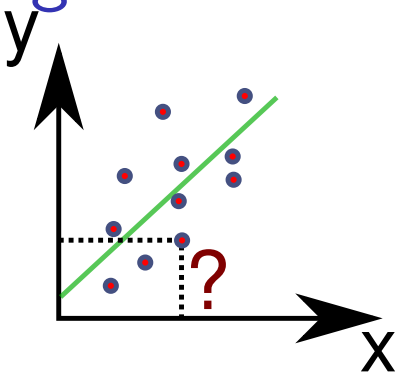


statsmodels - libreria per l'analisi statistica *classica*

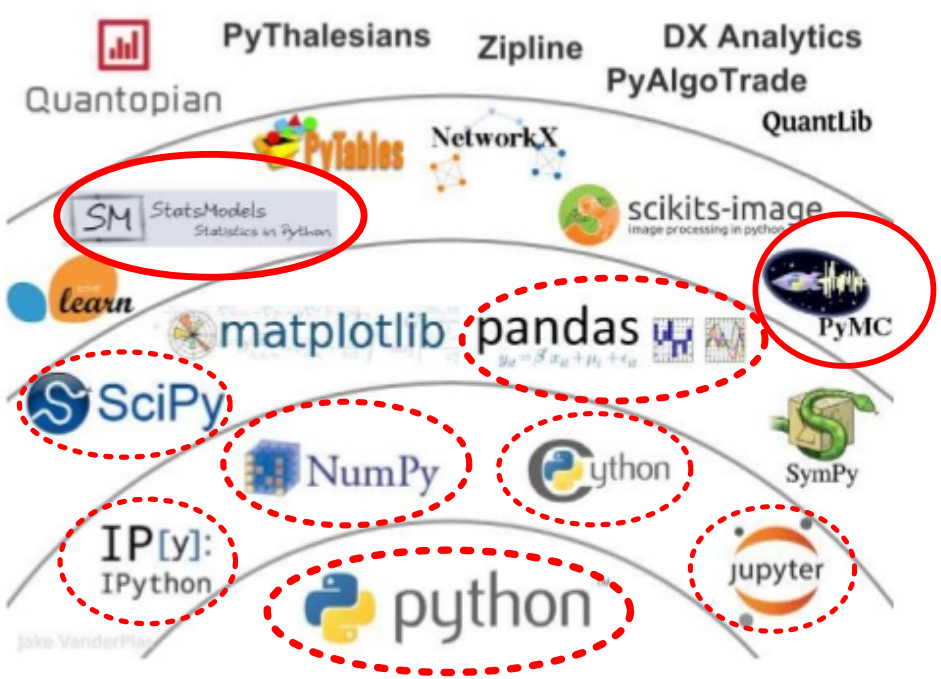


pymc2, **pystan** - librerie per l'analisi statistica *bayesiana*

Spiegare

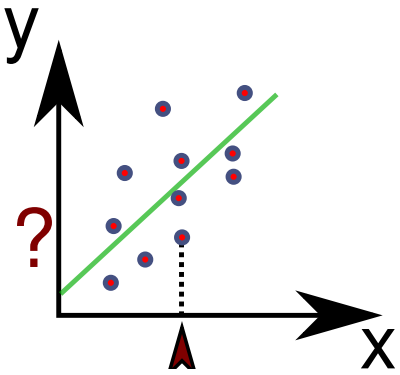


statsmodels - libreria per l'analisi statistica *classica*

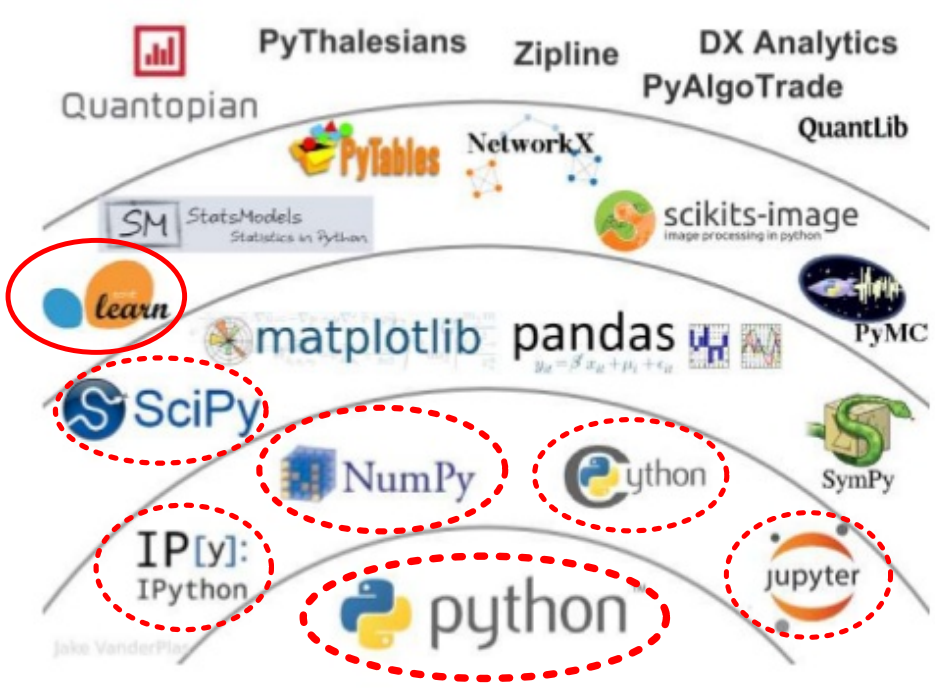


pymc2, **pystan** - librerie per l'analisi statistica *bayesiana* **Codice!**

Predire

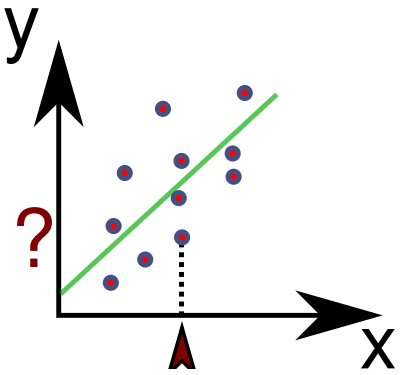


scikit-learn - libreria per machine learning

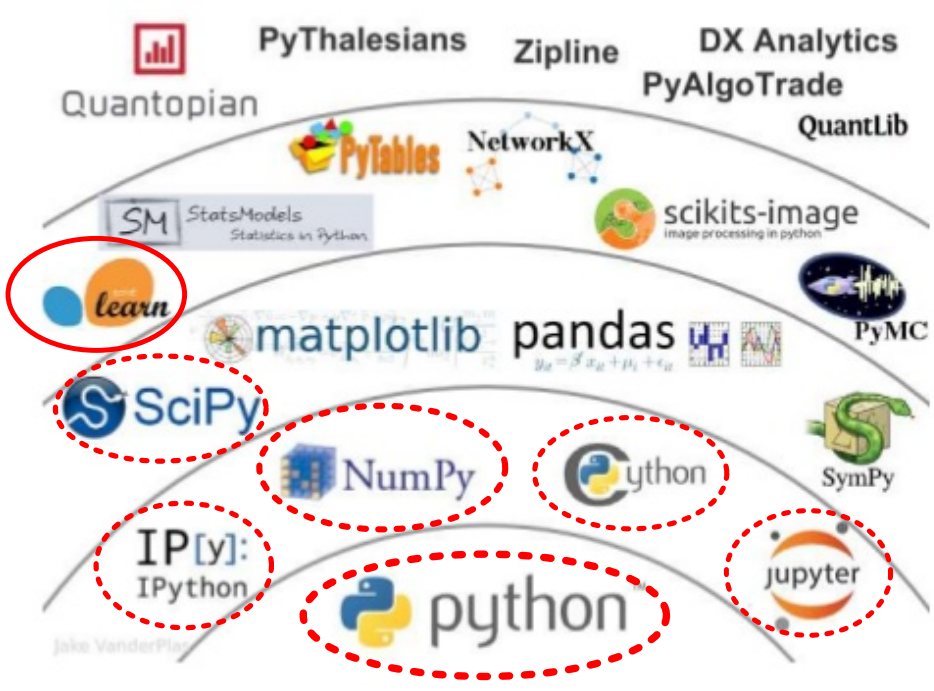


- Source: [Jake VanderPlas: State of the Tools] (https://www.youtube.com/watch?v=5GIlNDD7qbP4)

Predire



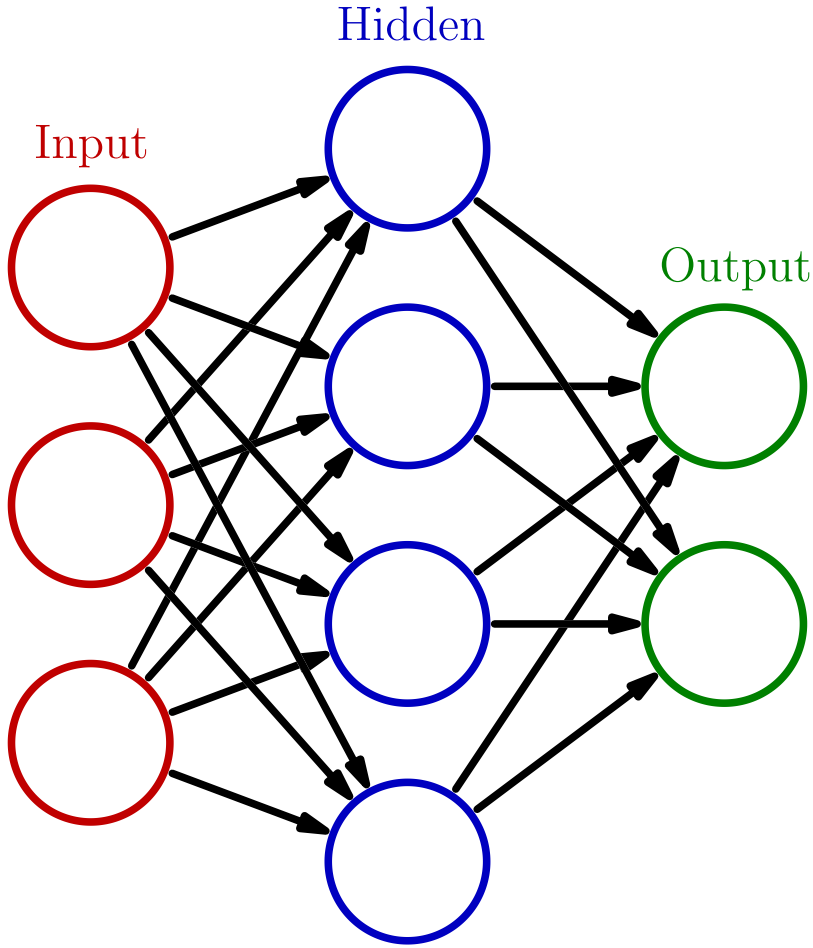
scikit-learn - libreria per machine learning



- Source: [Jake VanderPlas: State of the Tools] (https://www.youtube.com/watch?v=5GIlNDD7qbP4)

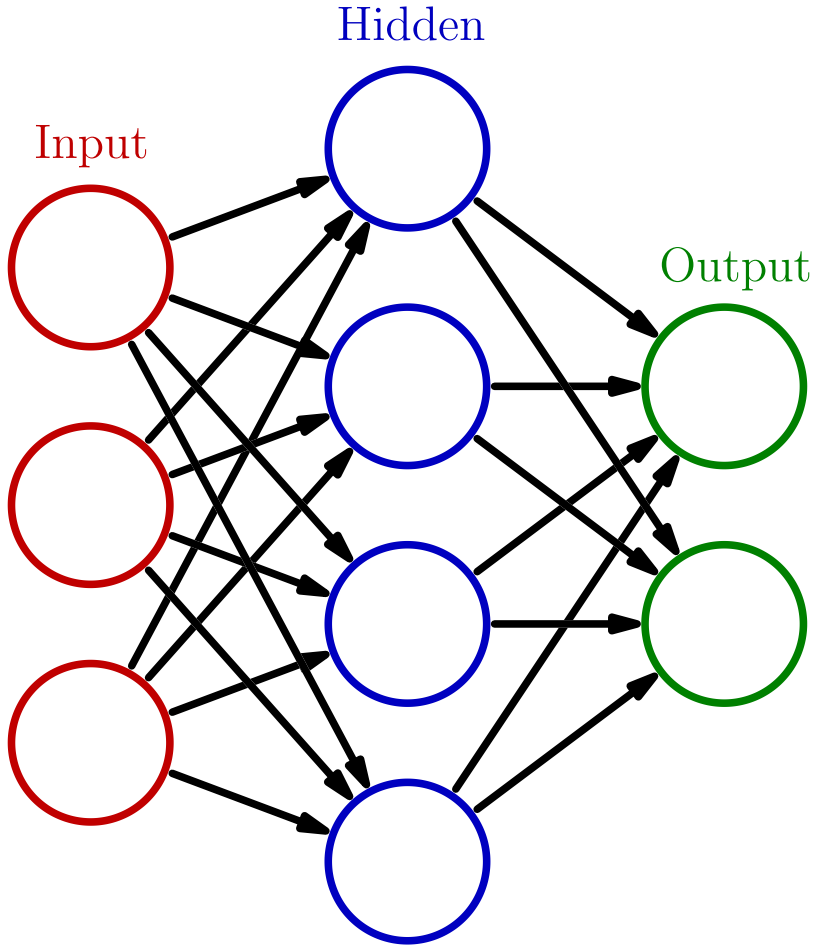
Codice!

Predire - con le reti neurali



Tensorflow, Theano, Keras, Caffe, Torch... - specifiche per neural networks

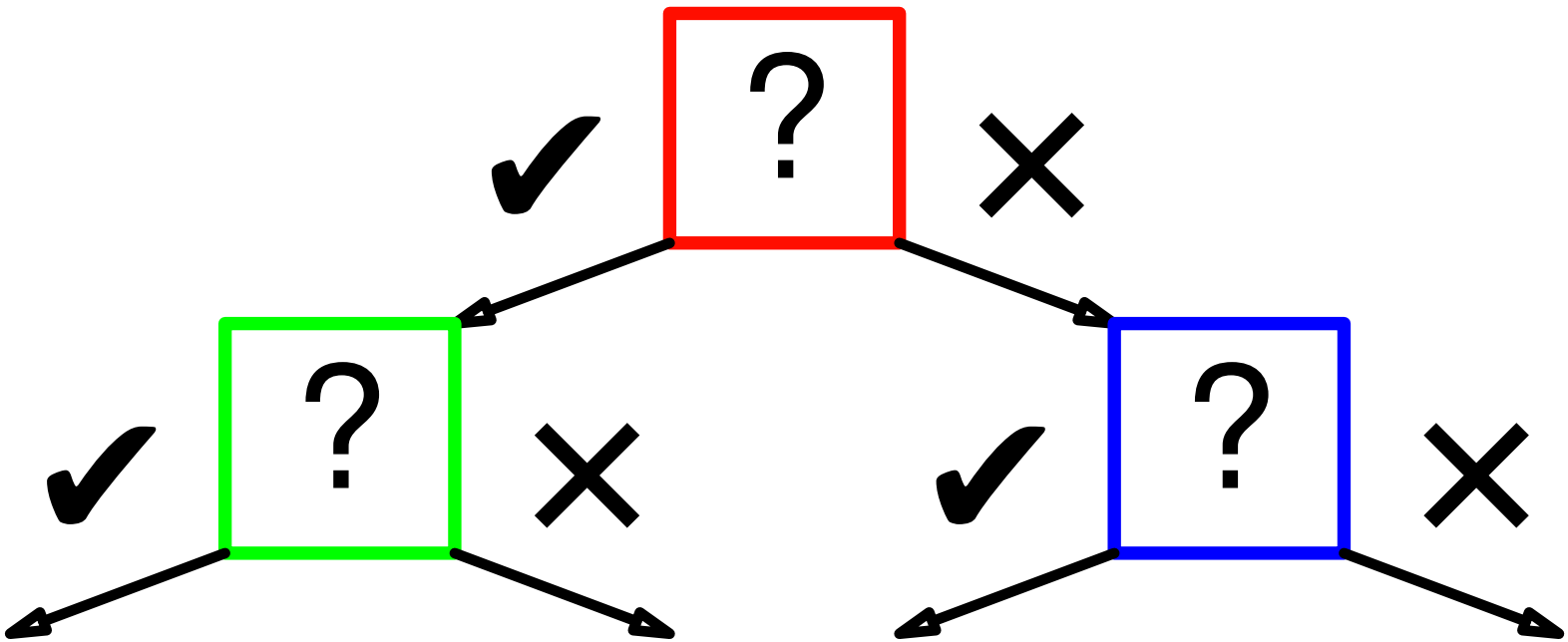
Predire - con le reti neurali



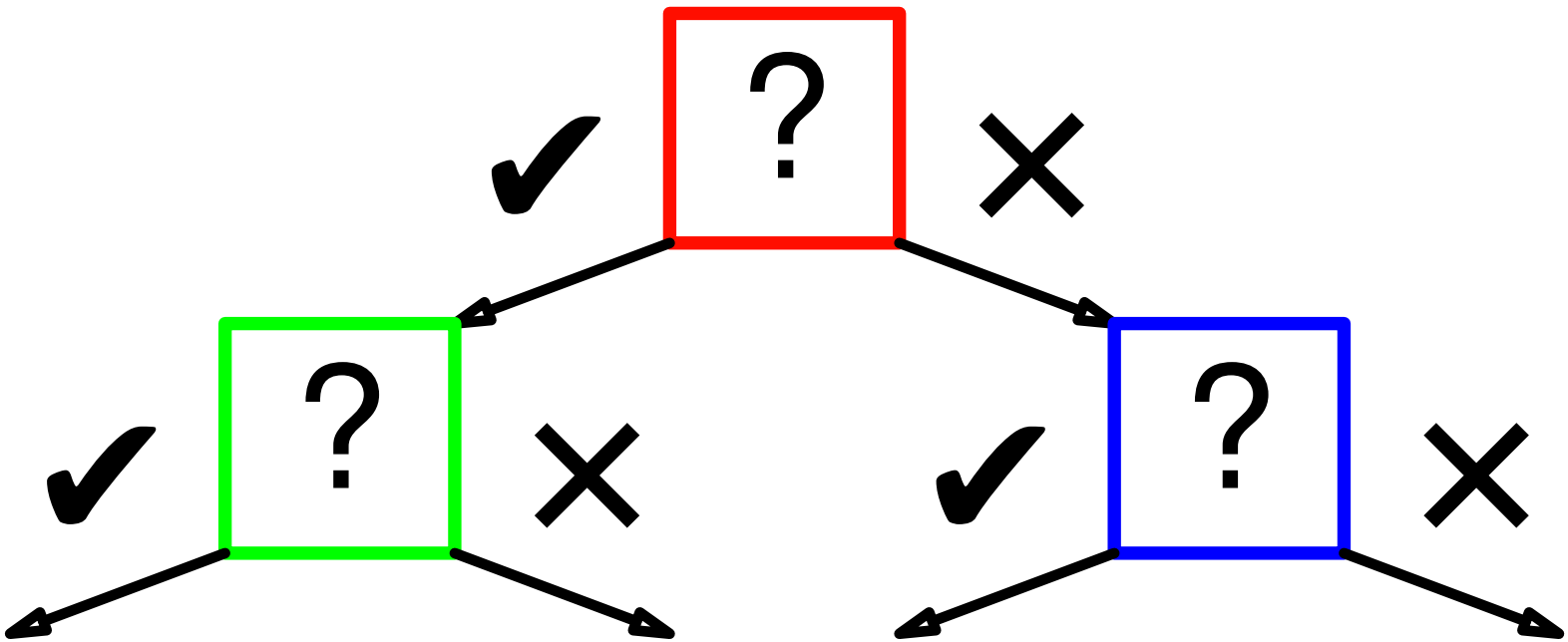
Tensorflow, Theano, Keras, Caffe, Torch... - specifiche per neural networks

Codice!

Predire - con un albero decisionale



Predire - con un albero decisionale



Codice!

Python o R?

Entrambi - rpy2

Credits

- Mappa dello stack Python per data analysis: **Jessica Stauth**: `https://www.slideshare.net/JessStauth/pydata-nyc-2015`
- **Wikipedia** per l'immagine del neural network: `https://en.wikipedia.org/wiki/File:Colored_neural_network.svg`