Hillsboro Python Machine Learning Meetup

Sep/2017

Ernest Bonat, Ph.D.

Senior Software Engineer

Senior Data Scientist

DAT Wi-Fi

Username: DAT Guest

Password: beaverton dat

- 6:00 – 6:40 pm: Pizza, **water only** and networking.
- 6:40 – 6:45 pm: Welcome message by Ernest Bonat, Ph.D.
- 6:45 – 8:00 pm: Presentation and open discussions.
- 8.00 pm – 9.00 pm: Coding and learning session. Bring your Python development laptop!

Why did I create this meetup?

1. Bad traffic to Portland downtown.

2. Vert hard to find a parking lot.

3. Bad Python presentation code.

4. No time at all to review the presentation and learn something after the meeting.

We need your support:

1. Need 1 Senior Python Developers for presentation and code review every month (Co-organizers, 4-6 hours a month).

2. Email Ernest at ebonat@15itresources.com

Our Meetup Mission:

1. *"Come, Listen, Code and Learn".*

2. Finding and presenting best practices of Machine Learning using Python Data Stack.

3. Create great networking place for Hillsboro-Beaverton Data Scientists.

"High Performance Big Data Analysis Using NumPy, Numba and Python Asynchronous Programming"

Dataconomy media (http://dataconomy.com/).

Here is the link: http://dataconomy.com/2017/07/big-data-numpy-numba-python/

Today Presentation

"Using Random Forests for Data Classification. Refactoring Python Data Science Projects"

Decision Trees main issue: **Overfitting** (the learning algorithm continues to develop hypotheses that **reduce training set error at the cost of an increased test set error**)

Random Forest model is a collection of **Decision Tree** models that are combined together to make predictions.

Random Forest is like bootstrapping algorithm with Classification and Regression Decision Treed (CART). Random forest tries to build multiple CART model with different sample and different initial variables.

Refactoring Python Data Projects

Tree Main Programming Styles:

1. **Top-bottom code** (like Jupyter Notebook) – teaching and presentations.

2. **Procedures code** – very old programing style.

3. **OOP code** – very necessary programming style for design, development and deployment enterprise software today.

Presentation Source Code

(https://github.com/ebonat/hillsboro_machine_learning_09_2017_2)