



Data Science and Machine Learning



- Python for Big Data Analytics and Machine Learning 101
- Deep learning with Tensorflow

Το Hellenic American Κε.Δι.Βι.Μ.2 προσφέρει σε συνεργασία με την εταιρεία Ismion Inc – MLTrain μια νέα σειρά μαθημάτων και εργαστηρίων στην επιστήμη δεδομένων και την μηχανική μάθηση (data science and machine learning), δύο κλάδους της πληροφορικής με μεγάλη δημοτικότητα τα τελευταία χρόνια. Η επιστήμη δεδομένων περιλαμβάνει το big data management, την σημαντικότερη ίσως νέα τεχνολογία στην περιοχή της πληροφορικής και η μηχανική μάθηση υπάρχει πλέον παντού: στα συστήματα αυτόματης διαπραγμάτευσης μετοχών, στ' αποτελέσματα αναζήτησης της Google, τ' αυτόνομα αυτοκίνητα, τους ψηφιακούς βοηθούς όπως οι Cortana και Siri και στα καταστήματα-χωρίς-ταμεία της Amazon. Η μηχανική μάθηση έχει επίσης χρησιμοποιηθεί με επιτυχία από CRM business units για την αυτόματη κατηγοριοποίηση της πελατειακής βάσης μίας εταιρίας και την δημιουργία έξυπνων καμπανιών προσέλκυσης και διατήρησης πελατών.



Μασσαλίας 22, 10680 Αθήνα

T: 210 3680966, 210 3680912 e-mail: it@haec.gr <http://www.haec.gr>

Python for Big Data Analytics and Machine Learning 101

Το σεμινάριο διάρκειας 16 ωρών, είναι μία επισκόπηση των βασικών μεθόδων και πρακτικών επεξεργασίας δεδομένων και μηχανικής μάθησης σε Python. Ο κώδικας που θα διδαχθεί τρέχει χωρίς διαφοροποίηση στο laptop και σε big data clusters. Οι συμμετέχοντες θα μάθουν ειδικότερα πώς να επεξεργάζονται δεδομένα χρησιμοποιώντας τις βιβλιοθήκες numpy και pandas, πώς να χρησιμοποιούν την PySpark για μεγάλα big-data clusters, πώς να δημιουργούν και να επικοινωνούν αντιπροσωπευτικές γραφικές απεικονίσεις με τις βιβλιοθήκες matplotlib και seaborn και πώς να εκπαιδεύουν μοντέλα μηχανικής μάθησης με το scikit-learn.

Διδακτέα Ύλη

1. Development front-ends: jupyter console, jupyter notebook and qtconsole

Using the command history
Interacting with the OS
The interactive debugger

2. Python bootcamp

Literals, expressions and statements
Python containers, comprehensions and generator expressions
Function objects, lambdas and closures

3. Fast array calculations with the numpy package

The ndarray object
Universal functions
Integer and boolean Slicing
Set logic

4. Tabular data management with the pandas package

Indexing, selection and filtering
Function application and mapping
Data filtering and reductions
Handling missing data
Hierarchical indexing

5. Cluster computing with Spark and PySpark

Installing and configuring Spark over Spark's standalone cluster
pyspark.Dataframes and untyped operations
running SQL programmatically
schema objects and types
aggregations

6. Plotting with matplotlib and Seaborn packages

Matplotlib API primer
Figures, subplots, axes, lines and markers
Line and bar plots
Histograms and density plots
Scatter plots

7. Introduction to scikit-learn

Decision trees
Random forests
Gradient boosted trees

Deep learning with Tensorflow

Το Tensorflow είναι μία βιβλιοθήκη αριθμητικής βελτιστοποίησης ανοικτού κώδικα της Google. Από πλευράς συστήματος, το σημείο διαφοροποίησης του Tensorflow είναι η ικανότητα παραλληλισμού και διανομής υπολογισμών σε χιλιάδες CPUs and GPUs, γεγονός που το καθιστά κατάλληλο για χρήση σε μεγάλα clusters. Στο σεμινάριο διάρκειας 16 ωρών, μαθαίνουμε πώς να προγραμματίζουμε νευρωνικά δίκτυα χρησιμοποιώντας το Tensorflow, το σημαντικότερο σήμερα API ανοικτού κώδικα που η Google χρησιμοποιεί και στην δική της μηχανή αναζήτησης.

Διδακτέα Ύλη

1. Tensor Basics

The computational graph model - imperative vs declarative programming

Basic operations with tensors

Graph Inspection & Visualization with TensorBoard

2. Linear Algebra with TF

Operations on sparse and dense matrices and vectors

Kronecker Products in TF

From matrices to tensors

Tensor tiling: The map Operator

Reductions on tensors

Auto-differentiation: TF's distinguishing feature

3. Tensorflow I/O

Feeding data with placeholders

Exporting results

Splitting and processing inputs with minibatches

4. Numerical optimization In TensorFlow

Creating a symbolic objective function

Gradient computations: building a simple gradient-descent optimizer

Tweaking gradients: the compute_gradient and apply_gradient methods

Predefined TF optimizers: gradient descent, adagrad and adam optimizers

Predefined loss functions

Building a linear regression model with TF in 3 lines

5. Introduction to Neural Networks and deep learning

Fundamentals of Neural Nets

Stochastic gradient descent

The backpropagation algorithm

Building a neural network classifier

Learning embeddings of categorical variables

Διατίθενται χορηγικά πακέτα για εταιρείες.
Για πληροφορίες επικοινωνήστε με την MLTrain
(www.mltrain.cc)

Για περισσότερες πληροφορίες επικοινωνήστε μαζί μας ή επισκεφθείτε μας.