
CURSOS CLAD

Data mining utilizando Python

Jaime S. Cardoso – Faculdade de Engenharia da Universidade do Porto

26 de novembro de 2016, Departamento de Matemática da Universidade de Aveiro

Resumo. O curso é uma organização conjunta da CLAD (Associação Portuguesa de Classificação e Análise de Dados) e da Universidade de Aveiro (DEGEIT e DMAT) e propõe uma introdução aos vários tipos de tarefas de extração de conhecimento de dados (*Data Mining*), com ênfase na análise preditiva (tarefas de regressão e classificação de dados). Serão apresentados métodos/algoritmos representativos para cada tipo de tarefa e ilustrada a aplicação desses métodos, com recurso ao Python, a problemas reais selecionados. Será, ainda, feita uma breve introdução ao Python (para quem já sabe uma linguagem de programação) e apresentados e utilizados os principais pacotes para Data Mining em Python.

Destinatários. Todos os potenciais utilizadores de técnicas de *data mining/machine learning* (docentes, investigadores, alunos e profissionais de outra áreas) que necessitem de abordar problemas de análise de dados e de avaliar os resultados e compreender o funcionamento dos métodos mais adequados. Será introduzida a teoria básica, mas a ênfase será dada à aplicação dos conceitos com recurso ao *software* Python.

Duração e Calendarização. O curso terá a duração de 6 horas e 30 minutos, das 10:00 às 18:30. O programa mais detalhado encontra-se em anexo.

Condicionante. O curso funcionará com um número mínimo e máximo de 10 e 30 participantes, respetivamente. Inscrições não aceites serão remetidas para futuras edições do curso. Cada participante deverá trazer o seu computador portátil. Antecipadamente serão enviadas instruções para instalação do Python.

Investimento e Data limite de inscrição. O curso é gratuito para sócios CLAD com a quota de 2016 paga; o investimento para não-sócios CLAD é de 60€. A CLAD emitirá um certificado de participação. A data limite para inscrição é 11 de novembro de 2016.

Contacto. Caso esteja interessado em frequentar este curso, a ficha de inscrição em anexo deverá ser remetida para o seguinte endereço de e-mail: mail@clad.pt. O mesmo contacto deverá ser utilizado para quaisquer outros esclarecimentos.

Com os melhores cumprimentos,
Pel'A Direcção da CLAD
José Gonçalves Dias

CURSOS CLAD

Jaime S. Cardoso – Faculdade de Engenharia da Universidade do Porto
26 de novembro de 2016, Departamento de Matemática da Universidade de Aveiro

Programa do curso

Data mining utilizando Python

- 10:00 – 13:00: O ABC do processo de aprendizagem: revisão dos conceitos principais do processo de aprendizagem automática/*data mining* com um foco na tarefa de classificação e sob o pressuposto da independência das observações.
- O ABC do Python e dos pacotes principais para *data mining*: comentário das características principais da linguagem Python e discussão dos pacotes NumPy e Scikit-learn.
- 14:30 – 16:30: Aplicação dos conceitos previamente discutidos em Python, propondo aos participantes a resolução de um de três problemas reais à escolha.
- 17:00 – 18:30: Análise de dados dependentes do contexto, em particular dados sequenciais. Motivação e discussão de Modelos de Markov Escondidos (HMM). Exemplificação/aplicação em Python.
-