

## **BASES DE DADOS AVANÇADAS**

Mestrado em Engenharia Informática

---

Código: 28501

Área Científica Predominante: Sistemas de Informação e Inteligência Artificial

Docente: Joaquim Gonçalves Pereira da Silva

Idioma de Instrução: Português

Regime: S1

Carga Letiva: 60h Carga Trabalho: 100h

ECTS: 6,0

---

### **Objetivos**

A Unidade Curricular (UC) de Bases de Dados Avançadas (BDA) pretende consolidar os conhecimentos de base de dados relacionais e introduzir os alunos a outros paradigmas de bases de dados. Esta UC irá incluir ainda tópicos sobre análise e o processamento de dados assentes em big data e em plataformas na nuvem, nas vertentes operacional e analítica.

### **Resultados da Aprendizagem**

No final desta unidade curricular, os alunos deverão ser capazes de:

1. Desenvolver soluções de manuseamento de dados, eficientes e seguras, baseadas em sistemas de gestão de bases de dados relacionais;
2. Compreender os modelos e os sistemas de bases de dados orientados a documentos, chave/valor, colunas e grafos;
3. Modelar e implementar soluções sobre os sistemas de gestão de bases de dados adequados ao tipo de dados e requisitos funcionais;
4. Desenvolver soluções eficientes para o processamento de grandes volumes de dados em termos operacionais e analíticos.

### **Conteúdos Programáticos**

1. Bases de dados relacionais
  - Sistemas de gestão de bases de dados
  - Triggers, procedimentos e funções
  - Índices e particionamento
  - Otimização de bases de dados
2. Bases de dados NoSQL
  - Formatos de dados NoSQL
  - Tipos de bases de dados NoSQL
  - Modelação de dados NoSQL
  - Exemplos práticos em MongoDB e Neo4j
3. Big data
  - Introdução a Big Data;
  - Tipos de ficheiros Big Data;
  - Sistemas distribuídos de ficheiros (Hadoop HDFS);
  - Arquitetura de sistemas Big Data: objetivo e tecnologias;
  - Arquitetura para data lake: data lake lake house, data mesh;

- Hive Data Warehouse: estrutura, manipulação e casos práticos;
- Processamento de dados: MapReduce, Spar.

#### 4. Big Data Analytics

- Carregamento de dados para Hive
- Extração, transformação e análise de dados com Spark

#### **Bibliografia Recomendada**

Kleppmann, M. (2020). Designing data-intensive applications: The big ideas behind reliable, scalable, and maintainable systems. O'Reilly Media, Inc. ISBN 978-1449373320

Luu, H. (2018). Beginning apache spark 2: with resilient distributed datasets, spark SQL, structured streaming and spark machine learning library. Apress. ISBN 978-1484235805

Sadalage, P. J., Fowler, M. (2012). NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence. Addison-Wesley. ISBN 978-0321826626

Sharda, R., Delen, D., Turban, E. (2017). Business Intelligence, Analytics, and Data Science: A Managerial Perspective. 4th edition, Pearson. ISBN 978-0134633282

Shaw, S., Vermeulen, A. F., Gupta, A., Kjerrumgaard, D. (2016). Practical Hive: A Guide to Hadoop's Data Warehouse System. Apress. ISBN 978-1484202722

Silberschatz, A., Korth, H. F., Sudarshan, S. (2019). Database System Concepts, 7th Ed. McGraw-Hill Education. ISBN 978-0078022159

#### **Métodos de Ensino e de Aprendizagem**

O programa cobre o âmbito de todos os objetivos da unidade curricular, complementado os conteúdos abordados com uma forte componente prática de implementação e teste. A realização de exercícios nas aulas e a elaboração de projetos em equipa ao longo do período de aulas permitirá consolidar os conhecimentos teóricos e dotar os alunos de competências técnicas muito valorizadas no mercado de trabalho.

#### **Métodos de Avaliação**

Serão realizadas aulas teóricas e outras fortemente práticas. Nas primeiras serão apresentados conceitos, abordagens, metodologias e tecnologias fundamentais. Nas aulas práticas serão resolvidos problemas tipo com soluções propostas, promovendo-se uma discussão participada dos alunos. A avaliação é composta pela realização de dois projetos práticos.

A nota final (NF) é a média pesada calculada segundo a expressão seguinte:

$$NF = \text{Nota projeto 01 (P01)} * 50\% + \text{Nota Projeto 02 (P02)} * 50\%$$

A avaliação inclui uma sessão de defesa individual dos projetos, sendo a classificação atribuída individualmente a cada elemento do grupo de trabalho. O aproveitamento à unidade curricular está sujeito à obtenção da nota mínima de 9,0 valores (escala de 0 a 20) a cada um dos projetos (P01  $\geq$  9 P02  $\geq$  9). A avaliação na época de exames consiste na entrega de novo projeto e na apresentação e defesa dos mesmo.