



UNIVERSIDADE FEDERAL DE SANTA CATARINA – UFSC
CENTRO DE CIÊNCIAS DA EDUCAÇÃO – CED
DEPARTAMENTO DE CIÊNCIA DA INFORMAÇÃO – CIN
CAMPUS UNIVERSITÁRIO - TRINDADE
CEP: 88040-970 - FLORIANÓPOLIS - SC
Fone: (048) 3721-4075 E-mail: cin@contato.ufsc.br

PROGRAMA DE ENSINO

1 IDENTIFICAÇÃO

Disciplina: **CIN7915 Data Science**

Carga Horária: 72 h/a – 4 créditos (2 teóricos, 2 práticos)

Oferta: Optativa para os Cursos de Graduação em Ciência da Informação, Biblioteconomia, Arquivologia e áreas afins.

EMENTA

Extração de informação a partir de dados. Definição de domínios de interesse. Segurança de dados. Mineração de dados científicos. Data Warehouse.

2 OBJETIVOS

2.1 Objetivo Geral

Conhecer os fundamentos sobre ciência de dados, assim como a utilização dos dados como fator de vantagem competitiva.

2.2 Objetivos Específicos

2.2.1 Conhecer os tipos de dados;

2.2.2 Entender os fundamentos de análise exploratória e preditiva;

2.2.3 Estudar os fundamentos da ciência de dados;

2.2.4 Aplicar os principais métodos e ferramentas para análise e ciência de dados.

3 CONTEÚDO PROGRAMÁTICO

3.1 Tipos de Dados

3.1.1 Dados estruturados

3.1.2 Dados semiestruturados

3.1.3 Dados não-estruturados

3.2 Armazenamento dos Dados

3.2.1 Modelos de armazenamento

3.2.2 SQL, NoSQL e NewSQL

3.2.3 Ferramentas para armazenamento de dados.

3.3 Filtragem e Modelagem de Dados

3.3.1 Análise de conteúdo

3.3.2 Estruturação e padronização de dados

3.3.3 Deduplicação de dados

3.3.4 Enriquecimento e modelos de dados

3.4 Métodos, Técnicas e Ferramentas

3.4.1 Mineração de dados

3.4.2 Data Warehouse

3.4.3 ETL, OLAP e OLTP

3.4.4 Agrupamento e classificação de dados

3.4.5 MapReduce, HDFS e Hadoop

3.5 Introdução à Segurança dos Dados

3.5.1 Aspectos gerais

3.5.2 Métodos computacionais para segurança de dados

3.5.3 Ferramentas computacionais para segurança de dados

4 BIBLIOGRAFIAS

4.1 Bibliografia Básica

AMARAL, Fernando. **Introdução à Ciência de Dados: Mineração de Dados e Big Data**. Rio de Janeiro: Alta Books, 2016. 320p.

DA SILVA, Leandro; PERES, Sarajane; BOSCARIOLI, Clodis. **Introdução à Mineração de Dados com Aplicações em R**. Rio de Janeiro: Elsevier, 2016. 277p.

DE CASTRO, Leandro; FERRARI, Daniel. **Introdução à Mineração de Dados: Conceitos Básicos, Algoritmos e Aplicações**. São Paulo: Saraiva, 2016. 351p.

PROVOST, Foster; FAWCETT, Tom. **Data science for business**. Sebastopol: O'Reilly, c2013. xxi, 386 p. ISBN 9781449361327.

4.2 Bibliografia Complementar

ATZENI, Paolo. **Database systems: concepts, languages & architectures**. London: McGraw Hill, c2000. xxiii, 612p. ISBN 007235872.

DATE, C. J. **Introdução a sistemas de bancos de dados**. Rio de Janeiro: Campus, 2004. 865p. ISBN 8535212736.

FINLAY, Steven. **Predictive analytics, data mining, and big data: myths, misconceptions and methods**. New York: Palgrave Macmillan, c2014. xii, 248 p. ISBN 9781137379276.

HURWITZ J. et al. **Big Data for Dummies**. New Jersey (USA): John Wiley & Sons, Inc., 2013. ISBN 978-1-118-50422-2.

MAYER-SCHÖNBERGER, Viktor; CUKIER, Kenneth. **Big data: a revolution that will transform how we live, work, and think**. Boston: Houghton Mifflin Company, c2013. 252 p. ISBN 9780544227750.

MILLER, Harvey J.; HAN, Jiawei. **Geographic data mining and knowledge discovery**. 2nd ed. Boca Raton: CRC Press, c2009 458 p. ISBN 9781420073973.

NISBET, Robert; ELDER, John; MINER, Gary. **Handbook of statistical analysis and data mining applications**. Amsterdam: Elsevier; London: Academic Press, 2009. xxxiv, 824 p. ISBN 9780123747655.

TAN, Pang-Ning; STEINBACH, Michael; KUMAR, Vipin. **Introdução ao datamining:** mineração de dados. Rio de Janeiro: Ciência Moderna, 2009. xxi, 900p. ISBN 9788573937619.