

3.7 Programa de Disciplina (Formulário SUPAC/UFBA)

UNIVERSIDADE FEDERAL DA BAHIA

SUPERINTENDÊNCIA ACADÊMICA PROGRAMA DE DISCIPLINAS SECRETARIA GERAL DOS
CURSOS**OBS: A carga horária básica deve ser de 17h ou múltiplo de 17h, conforme Res. 05/03 do CONSEPE.**

DISCIPLINAS

Código	Nome
	Fundamentos de Big Data

Carga Horária				Crédito	Assinatura do Chefe do Departamento	Ano
T	P	E	Total			
34	0	0	34			

Ementa / Objetivos

Ementa:

Introdução a Big Data, Métodos de amostragem em Big Data e Método de processamento em paralelo.

Objetivos:

Proporcionar ao aluno conhecimentos recentes em fundamentos e manuseio de grande volume de dados.

Metodologia

Conteúdo Programático

1. Introdução a Big Data
 - 1.1. Big Data: Por que e onde?
 - 1.2. Características de grandes volumes de dados
 - 1.3. Como e onde obter grandes volumes de dados
 - 1.4. Programação em Big Data – bibliotecas relevantes em R
 - 1.5. Aplicações práticas de Big Data

2. Métodos de amostragem em Big Data
 - 2.1. Bags of little Bootstrap
 - 2.2. Leveraging
 - 2.3. Log-verossimilhança média

3. Método de processamento em paralelo
 - 3.1. Estimação agregada (Aggregated Estimating Equations)
 - 3.2. Majority Voting

Bibliografia

Bibliografia básica:

1. Schöenberger, V. M. (2013). Big Data. Campus: Elsevier.
 2. Prajapati, V. (2013). Big Data Analytics with R and Hadoop.
 3. Wang, C et. Al. (2015). Statistical Methods and Computing for Big Data.
-
-