

BIG DATA



LÁB DATA



FUNDAÇÃO
INSTITUTO DE
ADMINISTRAÇÃO

Disciplina: Aplicações de Big Data com Hadoop

Tema da Aula: Aplicações MapReduce

Coordenação:

Prof. Dr. Adolpho Walter
Pimazzi Canton

Profa. Dra. Alessandra de
Ávila Montini

Profa. Rosangela de Fátima Pereira

Junho de 2016

Currículo

Formação

- Mestrado em Engenharia de Computação pela Escola Politécnica da Universidade de São Paulo (Poli-USP) (em andamento)
- Especialização em Tecnologia Java pela Universidade Tecnológica Federal do Paraná (UTFPR) (2011)
- Tecnologia em Análise e Desenvolvimento de Sistemas pela UTFPR (2011)
- Bacharelado em Administração de Empresas pela Universidade Estadual do Norte do Paraná (UENP) (2007)

Experiência

- Professora de Big Data Analytics em empresas e programas de MBA - FIA (2013 - atual)
- Pesquisadora no Laboratório de Arquitetura e Redes de Computadores (LARC) – USP (2013 - atual)
- Professora de cursos de engenharia na UTFPR (2011 -2012)
- Analista de sistemas na BSI Tecnologia (2009-2010)

LinkedIn: <https://br.linkedin.com/pub/rosangela-de-fatima-pereira/68/a10/b56>

Apaixonada por **Big Data!**

Objetivo da Aula

Implementar aplicações MapReduce para compreensão da biblioteca Java do Hadoop

Conteúdo da Aula

- Apresentação de projetos
- Aplicação 1 – Manipulação de dados no HDFS
- Aplicação 2 – Busca de top N Hashtags
- Desafio

Apresentação de projetos

Objetivo: compartilhar com todas as equipes as propostas de projetos de Big Data

Procedimento:

- 5 minutos de apresentação + 5 minutos de comentários

Aplicação 1

Manipulação de dados no HDFS

Objetivo: executar comandos do HDFS diretamente pela biblioteca JAVA do Hadoop.

Projeto base: HadoopHDFS

Aplicação 2

Busca de Top N Hashtags

Objetivo: permitir que o usuário envie como parâmetro da aplicação o valor N para busca de hashtags

Projeto base: Contador

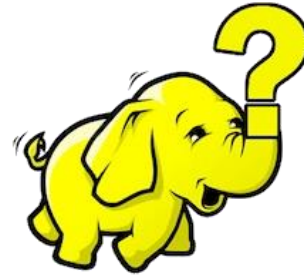
Desafio

1. Altere a classe TopNHashTagDriver para que seja verificado se o diretório informado pelo usuário já exista. Caso exista, exclua o diretório utilizando a API da biblioteca Java do Hadoop.
2. Desenvolva uma aplicação que faça a contagem de hashtags somente de uma determina palavra indicada pelo usuário como parâmetro.
3. Utilize a base de dados compras.txt e desenvolva uma aplicação que liste para cada cidade o tipo de produto mais vendido

Próxima aula

- Aplicação para recomendação de produtos aos usuários de e-commerce

Perguntas



rpereira@larc.usp.br

Referências Bibliográficas

WHITE, Tom. **Hadoop: The definitive guide**. " O'Reilly Media, Inc.", 2012.

PERERA, Srinath. **Hadoop MapReduce Cookbook**. Packt Publishing Ltd, 2013.