



Simpósio de Integração Acadêmica

“A Transversalidade da Ciência, Tecnologia e Inovações para o Planeta”
SIA UFV Virtual 2021



SENDAS: Extração de Padrões de Mobilidade de Grandes Volumes de Dados Geoespaciais

Universidade Federal de Viçosa, Instituto de Ciências Exatas e Tecnológicas – Campus Florestal

João Marcos A. M. Ramos, Fabrício A. Silva, Thais R. M. B. Silva

Palavras-chave: Dados geoespaciais, Fluxo de mobilidade, Ciência dos dados.

joao.m.ramos,fabricio.a.silva,thais.braga(@ufv.br)

Introdução

Neste trabalho, o objetivo é extrair conhecimento sobre a rotina coletiva de usuários móveis por meio do fluxo de deslocamento dos mesmos. Através da classificação de dados, é possível fazer agrupamentos dentro dos fluxos, observando diferentes comportamentos de acordo com a necessidade do usuário, como por exemplo, a separação entre horários do dia, ou dias da semana. A extração desses fluxos de deslocamentos é feita no contexto da biblioteca Sendas, desenvolvida no Núcleo de Estudos em Sistemas Pervasivos e Distribuídos (NESPED)

Objetivos

O objetivo geral do trabalho proposto em desenvolvimento tecnológico e inovação é a aplicação de técnicas de ciência dos dados para extrair conhecimento útil sobre a rotina de usuários móveis. Mais especificamente, o trabalho consiste em criar uma biblioteca para facilitar os estudos de dados geoespaciais, principalmente em relação à extração de fluxo de mobilidade.

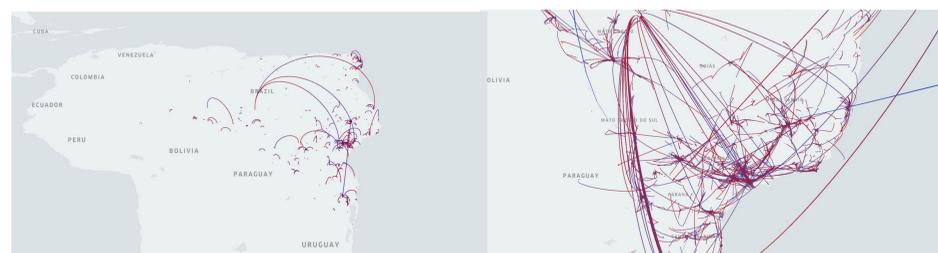
Material e Métodos

O primeiro passo antes de desenvolver o projeto foi a busca de uma linguagem e/ou framework que se adequaria com os objetivos, priorizando o desempenho e escalabilidade. Para isto, foi escolhido o framework Spark.

Em relação à linguagem, foi estudado Python e Scala, ambas apresentando vantagens e desvantagens, porém, foi escolhido Scala, devido ao seu ganho expressivo de desempenho.

Resultados e Discussão

Com a biblioteca, é possível extrair fluxos a partir de grandes volumes de dados geoespaciais, e visualizar esses fluxos de maneira clara, como podemos ver nas figuras abaixo.



Conclusões

Foi possível realizar a capacitação do bolsista, tanto na área de mineração de dados, quanto na análise estatística, junto com o aprendizado das ferramentas citadas nos materiais e métodos. O modelo proposto foi implementado e descrito, junto com as funcionalidades pensadas, e está em fase de testes para a integração com os produtos da empresa parceira do projeto.

Bibliografia

DAMJI, Jules S; WENIG, Brooke; DAS, Tathagata; LEE, Denny. Learning Spark. 2ª Edição. Local de publicação: O'Reilly Media, Inc, Julho de 2020.

Kon, F., Ferreira, É.C., de Souza, H.A. et al. Abstracting mobility flows from bike-sharing systems. Public Transp (2021).

Guo, Diansheng & Zhu, X. & Jin, Hai & Gao, Peng & Andris, Clio. (2012). Discovering Spatial Patterns in Origin - Destination Mobility Data.

Apoio Financeiro

CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico



Agradecimentos

