Saúde Pública

01530511 - Caio Victor Souza de Moraes

01560971 - Eduardo Aleixo Melo Gama

01157503 - Nilson Rodrigues de Mendonça Filho

01559154 - Renata Cristina da Silva Sousa

01542067 - Roberval Gonçalves de Farias Júnior

Olinda, 2024

**1. Introdução**

Este projeto busca investigar questões fundamentais de saúde pública, com foco na análise de dados para identificar padrões, tendências e desafios relacionados ao cuidado de crianças autista. Utilizando técnicas avançadas de Big Data, aprendizado de máquina e ferramentas estatísticas, serão analisados dados provenientes de sistemas de saúde.

**2. Objetivos**

* Coletar e organizar dados de sistemas de saúde sobre o cuidado de crianças autistas.
* Aplicar técnicas de Big Data e aprendizado de máquina para identificar padrões e tendências em questões de saúde pública voltadas para essa população.
* Analisar os desafios e barreiras no acesso a cuidados adequados e identificar áreas de necessidade prioritária para políticas públicas.
* Visualizar os dados analisados para fornecer insights e apoiar a formulação de políticas mais eficazes e inclusivas voltadas para o cuidado de crianças autistas.

**3. Arquitetura do Sistema**

A arquitetura do projeto segue o padrão de pipelines de Big Data e análise preditiva:

* **Coleta de Dados**: Dados foram coletados dos sistemas de saúde, e dados demográficos. Site de sistemas de saúde foi utilizado para extrair informações relevantes sobre o cuidado de crianças no espectro autista.
* **Processamento de Dados**: Utilizado Relatório do TEA para analise e criação de uma planilha em Excel para processar e limpar os dados, combinando diversas fontes para criar um conjunto de dados coeso. Algoritmos de aprendizado de máquina e análise estatística foram aplicados para identificar padrões de saúde pública, incluindo barreiras de acesso a cuidados e áreas críticas de necessidade.
* **Armazenamento**: Os dados foram armazenados em uma planilha no formato do excel distribuído para garantir escalabilidade e consultas rápidas, facilitando o acesso e análise de grandes volumes de dados em tempo real.
* **Visualização**: Os resultados foram visualizados através de dashboards do Power BI, permitindo a visualização de padrões, tendências e insights sobre o cuidado de crianças autistas, apoiando a tomada de decisões e a formulação de políticas públicas mais eficazes.

**4. Tecnologias Utilizadas**

* **Big Data Framework:** Excel para processamento distribuído de grandes volumes de dados.
* **Armazenamento:** Armazenamento em local.
* **Coleta de Dados:** Site do sistema de saúde, relatório do TEA.
* **Visualização de Dados:** Ferramenta Power BI para criação de dashboards interativos e visualização de padrões.

**5. Coleta de Dados**

* **Fonte de Dados:** Dados foram coletados de sistemas de saúde, registros eletrônico. Além disso, dentro do relatório do TEA contêm dados demográficos do estado de Pernambuco do ano 2023.
* **Período de Coleta:** Os dados foram coletados ao longo de 2 semanas.
* **Volume de Dados:** Aproximadamente 500 registros foram coletados, combinando dados de saúde público.
* **Estratégia de Coleta:** Utilizamos técnicas de raspagem de dados e APIs para extrair informações relevantes sobre crianças autistas, barreiras de acesso a cuidados e padrões de tratamento.

**6. Processamento de Dados**

* **Limpeza dos Dados:** Dados duplicados, inconsistentes e irrelevantes foram removidos. Informações como tweets não relacionados ao tema, ruídos e dados incompletos foram filtrados. Também foi aplicada normalização de dados para padronizar formatos e categorias.
* **Integração de Dados:** As várias fontes de dados foram combinadas para criar um dataset robusto e coeso, utilizando algoritmos de fusão de dados para manter a integridade das informações.
* **Análise Preditiva:** Modelos de aprendizado de máquina foram treinados para identificar padrões no comportamento de cuidado de crianças autistas, prever picos de demanda por serviços de saúde e avaliar a eficácia das políticas públicas atuais.

**7. Armazenamento de Dados**

Os dados processados foram armazenados em uma planilha excel com capacidade de escalabilidade e desempenho em cenários diferentes. A estrutura da tabela foi definida para facilitar consultas rápidas e análises preditivas.

**8. Visualização dos Resultados**

Utilizamos o Power BI para criar dashboards interativos que permitem visualizar os dados de forma clara e intuitiva. Os dashboards mostram padrões e tendências relacionadas ao cuidado de crianças autistas, incluindo as principais barreiras no acesso a cuidados médicos.

* **Visualizações:**
  + Gráficos de calor mostrando as regiões com maior incidência de barreiras no acesso a cuidados.
  + Gráficos de linhas mostrando o número de diagnósticos de autismo ao longo do tempo.
  + Mapas geoespaciais mostrando as áreas com maior escassez de serviços especializados.
  + Gráficos comparativos dos números de crianças atendidas em diferentes tipos de centros de saúde.

**9. Resultados**

* **Incidência:** A análise dos dados identificou um aumento significativo no número de diagnósticos de autismo em regiões urbanas, enquanto áreas rurais enfrentam grandes dificuldades no acesso a diagnósticos e tratamentos especializados.
* **Barreiras Identificadas:** Barreiras como falta de profissionais especializados, longas filas de espera e custos elevados foram predominantes. Identificou-se que a falta de transporte e infraestrutura adequada também impacta negativamente o acesso a cuidados.
* **Eficácia das Políticas Públicas:** Embora algumas regiões tenham mostrado melhorias no acesso ao cuidado, há uma necessidade urgente de políticas públicas mais abrangentes, que priorizem a capacitação de profissionais de saúde e a expansão dos serviços especializados em áreas periféricas.

**10. Desafios**

* **Volume de Dados:** O grande volume de dados exigiu uma otimização cuidadosa do pipeline de analise em excel.
* **Qualidade dos Dados:** Houve muitos dados inconsistentes, especialmente provenientes de sites, exigindo uma filtragem rigorosa para garantir a relevância das informações.
* **Interoperabilidade:** A integração de dados de diferentes sistemas de saúde apresentou desafios de compatibilidade, exigindo conversões de formatos e ajustes nos padrões de interoperabilidade.

**11. Conclusão**

Este projeto demonstrou a eficácia do uso de técnicas de Big Data e aprendizado de máquina para identificar padrões e tendências no cuidado de crianças no espectro autista. As barreiras de acesso identificadas fornecem uma base sólida para a criação de políticas públicas mais inclusivas e focadas nas reais necessidades dessas populações. A visualização dos dados permitiu uma compreensão mais clara das áreas críticas de necessidade, promovendo ações governamentais mais direcionadas e eficazes.

12. **Referências**

* Ministério da Saúde (2023). Relatório Anual de Saúde Pública no Brasil.
* Relatório TEA – Transtorno do Espectro Autista.
* Informações de Saúde (TABNET) – DATASUS.