

Pontifícia Universidade Católica de Minas Gerais

Instituto de Ciências Exatas e Informática

Curso: Ciência da Computação

Disciplina: Inteligência Artificial

Professora: Cristiane Neri Nobre

28 de Fevereiro de 2019

Grupo

- Gabriel Luciano Gomes;
- Geovane Fonseca de Souza Santos;
- Luigi Domenico Cecchini Soares;
- Paulo Junio Reis Rodrigues.

Descrição da Base

Para o desenvolvimento do trabalho, será utilizado o conjunto de dados fornecidos pelo IBGE, na [Pesquisa Nacional de Saúde \(2013\)](#). A base apresenta informações relacionadas às condições da saúde da população brasileira, com enfoque em três perspectivas: percepção do estado de saúde, estilos de vida e doenças crônicas. Pensando nisso, os [dados](#) estão divididos em duas tabelas distintas, que possuem tanto atributos nominais quanto numéricos. São elas:

- Tabela Domicílios: contém cerca de 70 atributos relativos a moradia, apresentando questões como: tipo do domicílio (casa, apartamento), materiais predominantes nas construções do ambiente (piso, paredes, telhado), formas de abastecimento e disponibilidade de água, e bens existentes no recinto;
- Tabela Pessoas: é a tabela com maior volume de dados, contendo 205.546 instâncias e 942 atributos, distribuídos em módulos distintos, que procuram responder questões relacionadas à, por exemplo, condições de saúde da população brasileira (e.g. atividades físicas, doenças crônicas, deficiências), situações de violência pelas quais a pessoa passou (e.g. física, sexual, psicológica) e condições financeiras do entrevistado. Além disso, os dados apresentam como enfoque indicadores de saúde para grupos e faixas etárias distintas.

O Problema

A base de dados apresenta informações referentes as mais variadas condições físicas e mentais. Dentre elas, estão presente dados relativos a depressão e a deficiências físicas (e.g. física, auditiva, visual), sendo que cerca de 4000 pessoas apresentam quadro depressivo e cerca de 1100 apresentam algum tipo de deficiência. Existe, ainda, uma interseção entre esses dois tipos distintos de instâncias, com cerca de 600 pessoas possuindo depressão e algum tipo de deficiência. Pensando nisso, a ideia inicial era de responder ao seguinte questionamento: “Pessoas que apresentam algum

tipo de deficiência estão mais propensas a desenvolver quadros de depressão?”.

Após algumas análises e discussões realizadas, ficou definido que o questionamento deveria ser modificado de forma a abranger a quantidade total de instâncias que apresentam depressão. Dessa forma, a pergunta a ser respondida passa a ser: “Quais os perfis de pessoas com depressão?”. Assim, serão aplicadas técnicas de *clusterização*, para definir grupos com o maior grau de pureza possível e, com isso, tentar encontrar perfis distintos de indivíduos depressivos - entre eles, os que possuem algum tipo de deficiência.