

Interações entre Droga e Doença por meio de Genes

Mateus Siqueira Batista Nicolas Bissoli Nattis

MC536 - Instituto de Computação, UNICAMP

2020

Proposta

Obter dados de interações entre genes e drogas, e entre genes e doenças.

Proposta

Obter dados de interações entre genes e drogas, e entre genes e doenças.

Através disso, podemos relacionar a interação entre estas drogas e as doenças.

Proposta

Obter dados de interações entre genes e drogas, e entre genes e doenças.

Através disso, podemos relacionar a interação entre estas drogas e as doenças.

Exemplo:

- Droga A ativa o gene X.
- Gene X tem relação de causa com as doenças α , γ .
- Portanto, a droga A tem relação de causa com as doenças α , γ .



- Dados sobre interações droga-gene e o genoma drogável.
- Extraído de mais de trinta fontes confiáveis.
- Dados extraídos via arquivo .tsv disponibilizados.



- Plataforma contendo uma das maiores coleções publicamente disponíveis de genes e variantes associados a doenças humanas.
- Dados extraídos via banco de dados SQLite disponibilizado.

A relação de droga-doença ocorre de modo intuitivo a partir das relações de doença e droga com o gene.

		Droga-Gene	
		ativação	inibição
Gene-Doença	ativação	ativação	inibição
	inibição	inibição	ativação

Tipos de Interação

Activator

Activator

Inducer

Agonist

Binder

Chaperone

Partial Agonist

Potentiator

Stimulator

Substrate

Inhibitor

Antibody

Antagonist

Inhibitor

Antisense
Oligonucleotide

Blocker

Inhibitory
Allosteric
Modulator

Inverse Agonist

Partial Antagonist

Suppressor

DgiDb

Susceptibility
Mutation

Chromosomal
Rearrangement

Somatic Causal
Mutation

Fusion Gene

Germline Causal
Mutation

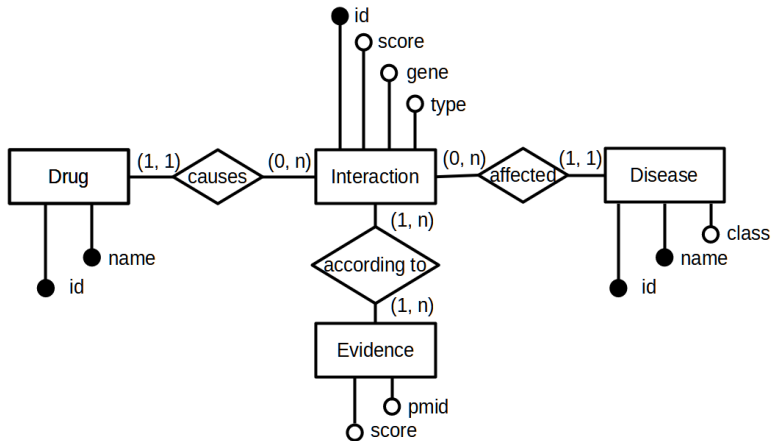
GeneticVariation

Biomarker

Therapeutic

DisgeNet

Modelo Conceitual



Modelo Lógico

Drug(DrugId, Name)

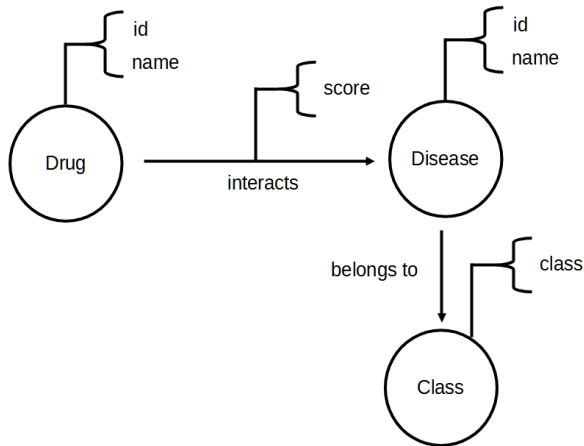
Disease(DiseaseId, Name, Class)

Modelo Lógico

```
Interaction(InteractionId,  
           DrugId,  
           DiseaseId,  
           Score,  
           Gene,  
           Type)
```

```
Evidence(InteractionId,  
         Pmid,  
         Score)
```

Modelo Lógico de Redes Complexas



Tratamento de Dados e Preparo do Dataset

- Quantidade muito grande de dados.
 - Mais de 3,2 milhões de interações gene/doença no DisgeNet.
 - Coleta pode se tornou lenta e tivemos que buscar alternativas de otimização.
- Confibialidade
 - Os dados coletados possuem níveis de confiabilidade variáveis, que foram levados em conta e tranformados em um Score.
 - Há um Score atrelado ao Dgldb e outro ao DisGeNET. Assim, obtivemos um Score global da interação multiplicando um pelo outro.

Perguntas de Análise

- Quais **drogas** tem relação com a **doença** Y?
- Quais classes de **doenças** estão mais relacionados com a **droga** Y?
- Quais são as evidências da relação entre uma **doença** X e uma **droga** Y?

Obrigado pela atenção!