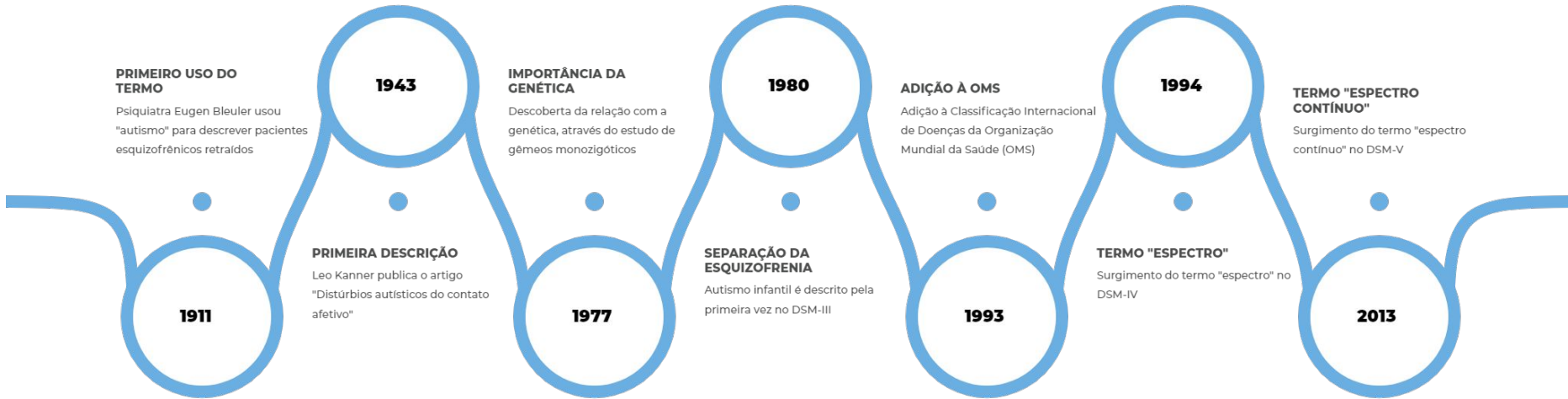

Diagnóstico do Transtorno do Espectro Autista (TEA)

Projeto Final

Ciência e Visualização de Dados em Saúde

Felipe Correia Labbate
Gabriela Servidone
Thiago Giachetto de Araujo

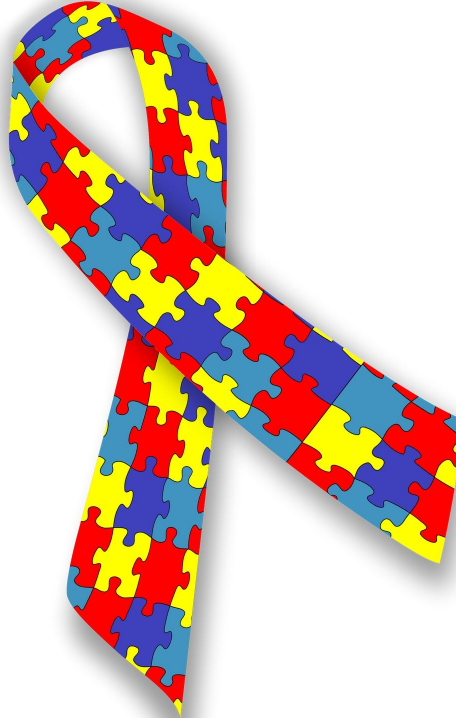
Autismo - Timeline



Diagnóstico do TEA

Diagnóstico clínico:
processo que envolve
diversos profissionais,
baseado no histórico e
observação do indivíduo

Sintomas se manifestam,
na maioria das vezes, até
os 3 anos de idade



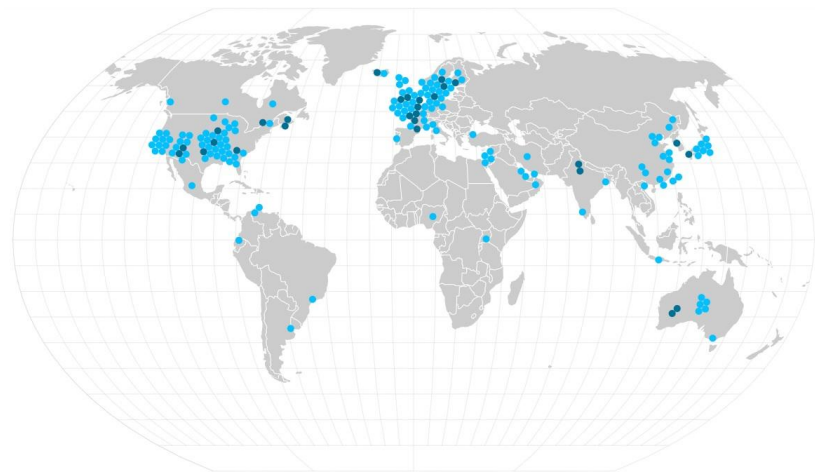
Diagnóstico precoce pode
ser complexo, e até
prejudicial quando feito de
forma errada

Intervenção precoce está
relacionada com ganhos
significativos no
desenvolvimento da criança

**Por isso o diagnóstico do Transtorno do Espectro
Autista (TEA) deve ser confiável e preciso!**



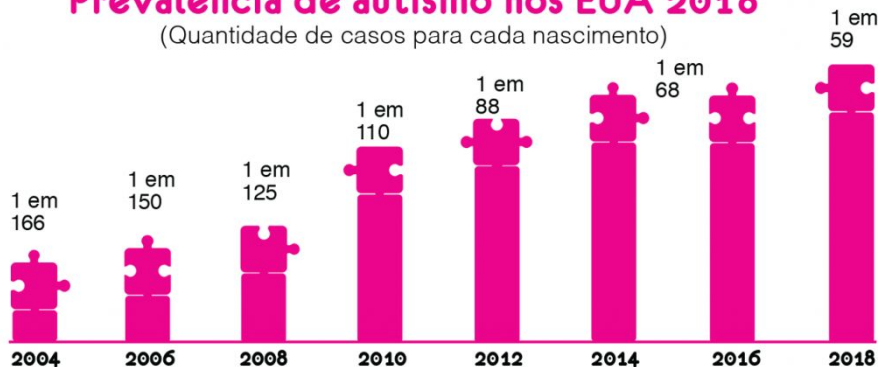
Prevalência do TEA - Pesquisas



Fonte: <https://prevalence.spectrumnews.org/>

Prevalência de autismo nos EUA 2018

(Quantidade de casos para cada nascimento)



Fonte: Centers for Disease Control and Prevention (CDC) - EUA



arte: Revista Autismo

Dados no Brasil e no mundo

Brasil atualmente **não possui** números oficiais sobre o autismo.

“Lei Romeo Mion” (Lei nº 13.977, de 2020) - Carteira de Identificação da Pessoa com Transtorno do Espectro Autista (Ciptea).

Dados globais

ONU (2013): 1% da população

OMS (2017): 1 em cada 160 = 0.625%

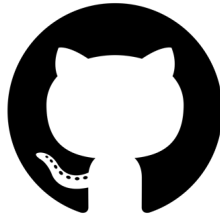
USA (2018): 1 em cada 59 \approx 1.7%

Ferramentas Utilizadas

colab

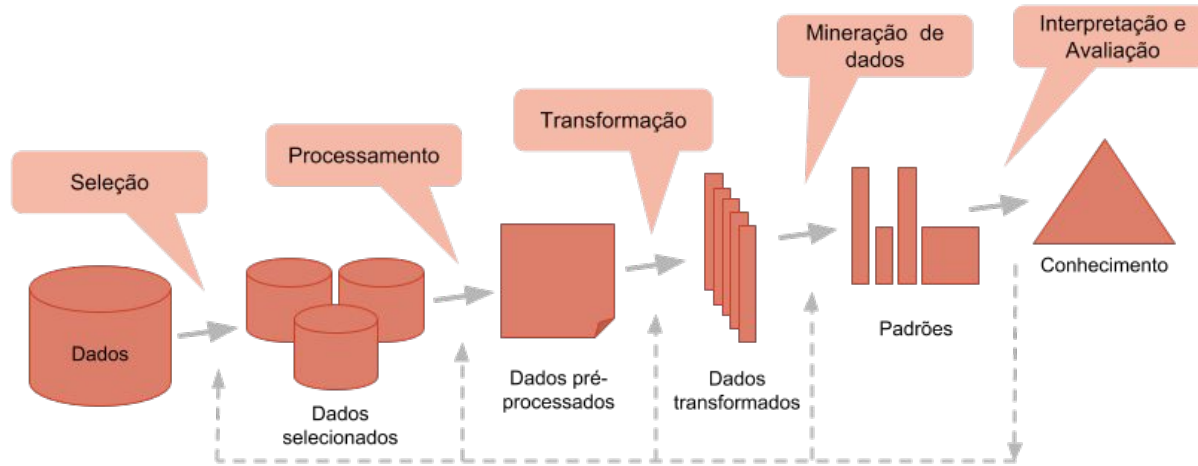


Google Drive



StackEdit

Modelagem



<https://medium.com/@kvmoura/kdd-process-9b8e3062142>

Knowledge Discovery in Database (KDD)

Base de dados e Seleção

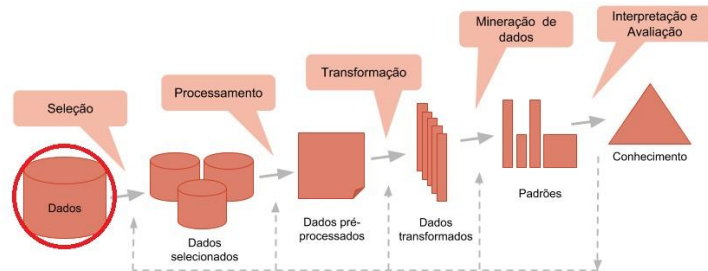
National Survey of Children's Health (NSCH) - survey anual da Data Resource Center for Child & Adolescent Health (DRC)

Perguntas gerais sobre aspectos da vida das crianças dos EUA, como saúde física e mental, histórico familiar, contexto social no qual está inserida, etc.

Dados sigilosos - formulário de aceite dos termos de sigilo

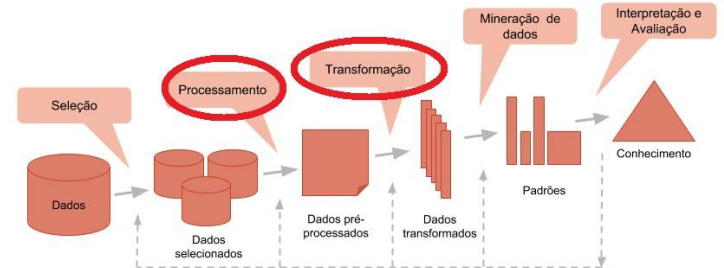
Dados correspondentes aos anos 2017-2018:

- 745 perguntas;
- 52.129 registros;
- 1.345 registros "Sim" na pergunta "Autism ASD Currently".



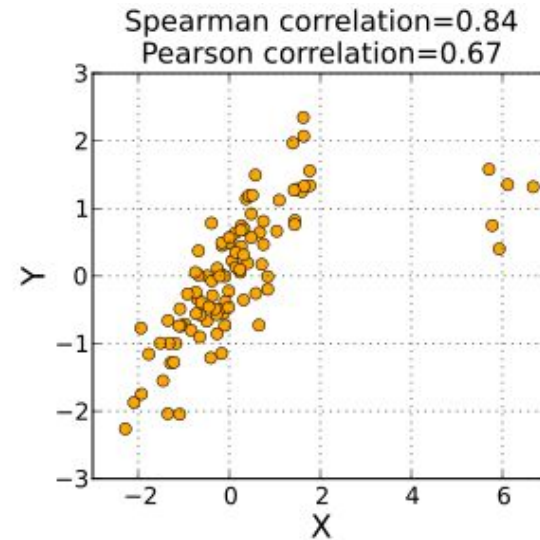
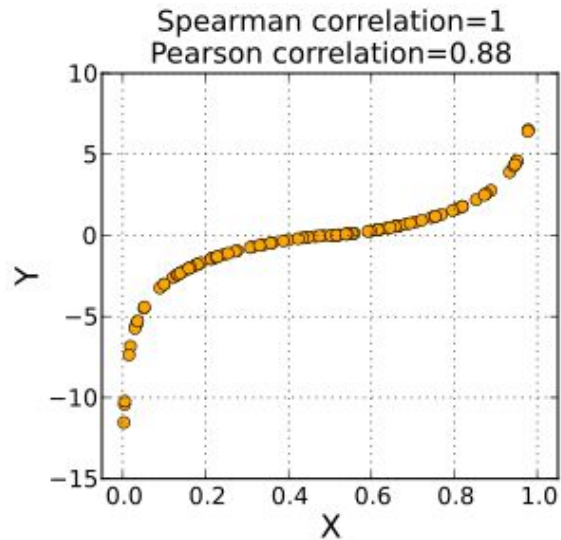
Pré-processamento e transformação

- Correlação de Spearman
- Filtragem de features
 - Diretamente relacionadas a TEA
 - Questões de screening
 - Correlação



Pré-processamento e transformação

Correlação de Spearman



Pré-processamento e transformação

Diretamente relacionadas a TEA

Questões de screening

asd

autism

sc

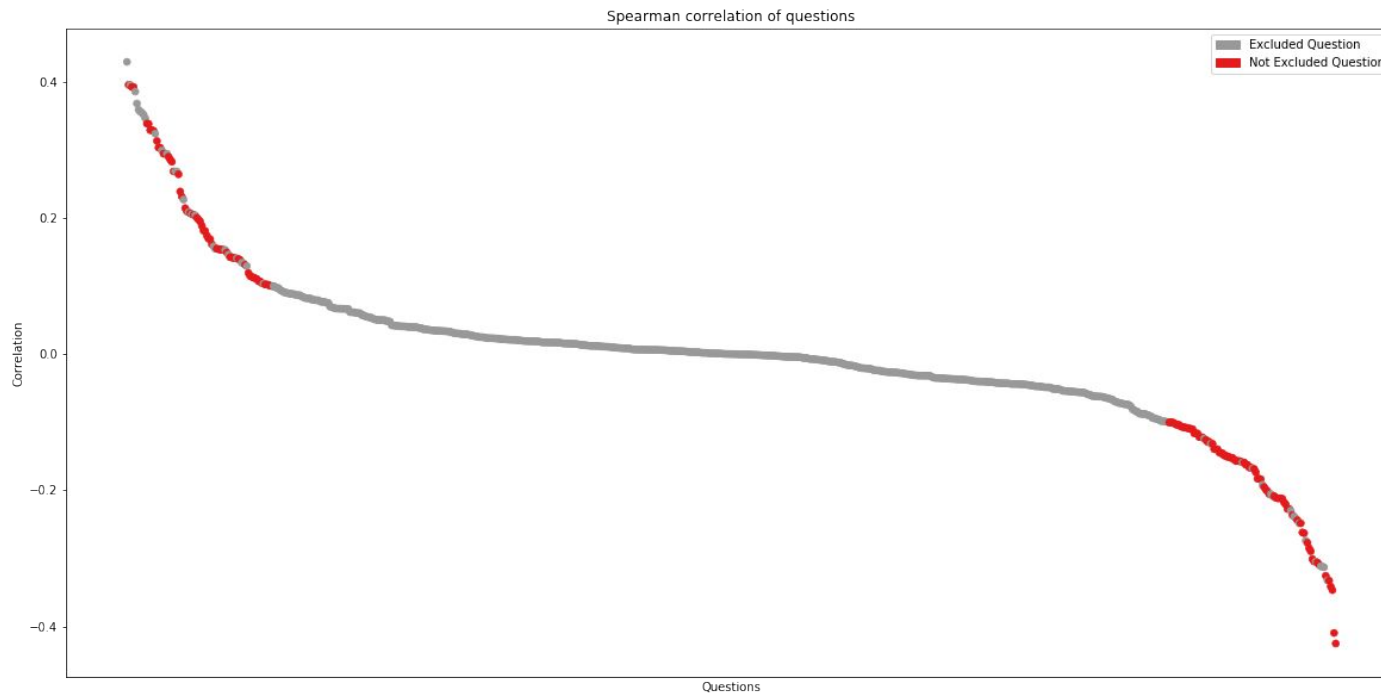
indicator

screener

CSHCn

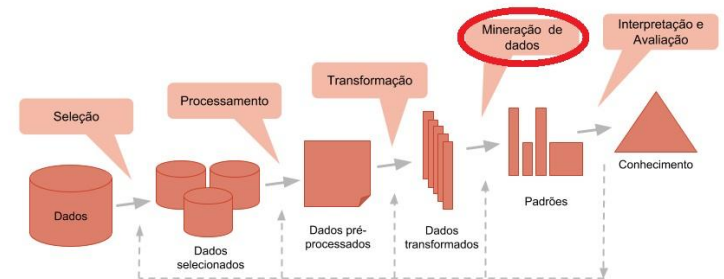
Pré-processamento e transformação

Filtragem pela correlação



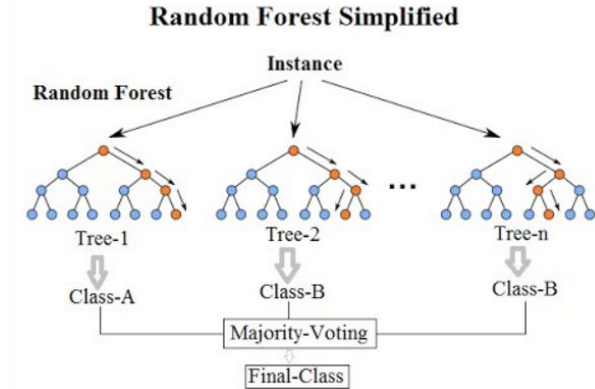
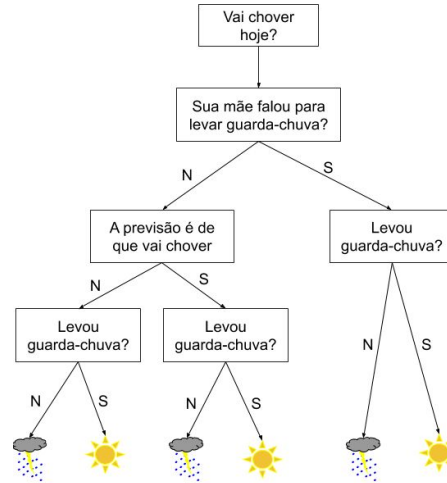
Mineração de dados

- **Seleção de features**
 - Árvore de decisão
 - Random Forest
- **Classificadores**
 - Support Vector Machine
 - Regressão Logística
 - Random Forest
- **Sampling**
 - Oversampling
 - Undersampling



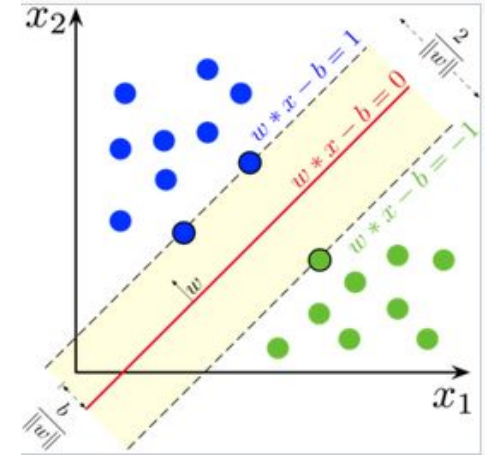
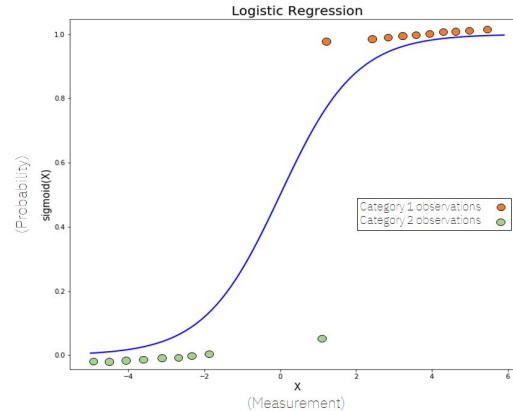
Mineração de dados

- **Seleção de features**
 - Árvore de decisão
 - Random Forest
- **Classificadores**
 - Support Vector Machine
 - Regressão Logística
 - Random Forest
- **Sampling**
 - Oversampling
 - Undersampling



Mineração de dados

- **Seleção de features**
 - Árvore de decisão
 - Random Forest
- **Classificadores**
 - Support Vector Machine
 - Regressão Logística
 - Random Forest
- **Sampling**
 - Oversampling
 - Undersampling

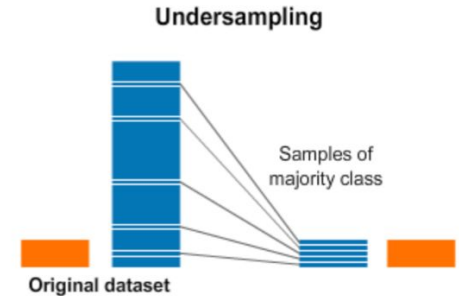
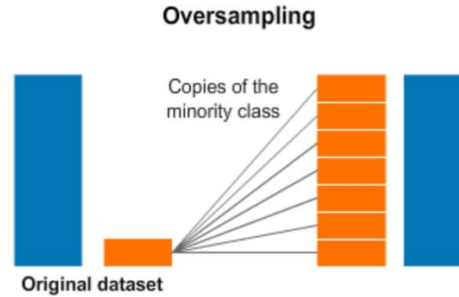


Fonte: https://en.wikipedia.org/wiki/Support_vector_machine

Fonte: <https://towardsdatascience.com/logistic-regression-explained-9ee73cede081>

Mineração de dados

- **Seleção de features**
 - Árvore de decisão
 - Random Forest
- **Classificadores**
 - Support Vector Machine
 - Regressão Logística
 - Random Forest
- **Sampling**
 - Oversampling
 - Undersampling



Mineração de dados

feature selection	classifier	precision*	recall*	f1-score*	Oversampling
Random Forest(***)	Support Vector Machine	0.82	0.62	0.67	N
Decision Tree Classifier(***)	Support Vector Machine	0.49	0.5	0.49	N
Decision Tree Classifier(**)	Logistic Regression	0.79	0.64	0.69	N
Decision Tree Classifier(***)	Logistic Regression	0.73	0.67	0.69	N
	Random Forest(****)	0.81	0.62	0.67	N
	Random Forest(****)	0.93	0.51	0.51	Y
Random Forest(***)	Support Vector Machine	0.81	0.62	0.67	Y

Interpretação e avaliação do resultado

ID	Variável	Descrição
1	ANYOTHER	Any Other Condition
2	CSHCNComplex_1718	Complexity of health care needs
3	MENBEVCOV	Health Insurance - Cover Mental Behavioral Needs
4	K4Q38	Received Special Services - Currently
5	ARRANGEHC	Hours Spent Arranging Health Medical Care
6	K5Q21	Arrange Or Coordinate Care Extra Help
7	TOTNONSHCN	Count of Children without Special Health Care Needs in House
8	partic_1718	Number of organized activities child participates in after school
9	EMOSUPADV	Emotional Support - Health Condition Support Group
10	SESPLANYSR	Special Education Plan - Age in Years
11	K4Q26	Specialist Visit - Problem
12	K2Q33C	Anxiety Severity Description

Modified Checklist for Autism in Toddlers

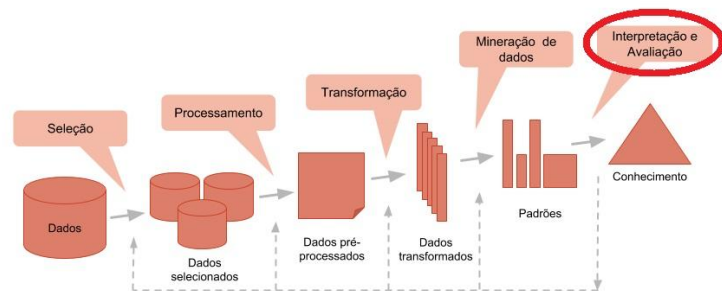
13	K5Q11	Need a Referral - Problem
14	K8Q32	How Often Have You Felt - Child Really Bothers You
15	K2Q01	General Health
16	C4Q04	Frustrated In Efforts to Get Service
17	K8Q21	Share Ideas or Talk About Things That Matter
18	K7Q85_R	Stays Calm and In Control When Challenged
19	cntdiff	Number of functional difficulties reported from list of 12 different difficulties asked in the survey
20	ADHDSev_1718	Parent-rated severity of child's current ADD/ADHD, age 3-17 years
21	K8Q31	How Often Have You Felt - Child Hard to Care For
22	HCEXTENT	Health Affected Ability - Extent
23	CondCnt27_1718	Number of current health conditions reported from a list of 27
24	speech_1718	Children who currently have speech or other language disorder, age 3-17 years
25	MAKEFRIEND	Difficulty Making or Keeping Friends

Interpretação e avaliação do resultado

É possível criar um classificador de pessoas autistas via perguntas específicas?

Existe correlação entre as respostas dos formulários e a classificação do paciente?

Como especialistas classificam se uma pessoa é ou não autista?



Dificuldades Enfrentadas

- Base de dados privada
- Especialistas em autismo

Mudanças de Percurso

- Cronograma

Lições Aprendidas

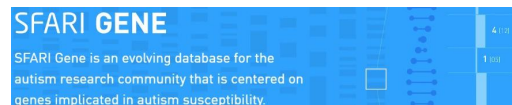
- Conhecimento teórico (TEA)
- Conhecimento técnico
 - KDD
 - Linguagens



Trabalhos futuros

Combinação: Perguntas comportamentais + Fatores genéticos

Exemplo de base de dados genéticos ligados ao autismo: SFARI Gene



STATUS	GENE SYMBOL	GENE NAME ↗	CHROMOSOME	GENETIC CATEGORY	GENE SCORE	SYNDROMIC	REPORTS
	ADSL	adenylosuccinate lyase	22	Rare Single Gene Mutation, Syndromic	1		6
	ADSS2	adenylosuccinate synthase 2	1	Rare Single Gene Mutation	3		3
	ADRB2	adrenergic, beta-2-, receptor, surface	5	Genetic Association	3		10
	AFF2	AF4/FMR2 family, member 2	X	Rare Single Gene Mutation, Syndromic	1		15
	ALDH1A3	aldehyde dehydrogenase 1 family member A3	15	Rare Single Gene Mutation, Syndromic	-	S	5
	ALDH5A1	aldehyde dehydrogenase 5 family, member A1 (succinate-semialdehyde dehydrogenase)	6	Rare Single Gene Mutation, Syndromic	1		10
	ALG6	ALG6, alpha-1,3-glucosyltransferase	1	Syndromic	-	S	1

Comentário da Psicóloga

"Este projeto apresentou de forma clara e objetiva o apanhado de informações que estão disponíveis na área de pesquisa a respeito do TEA, **dando atenção ao contexto biopsicossocial do transtorno**. Ainda é difícil realizar o diagnóstico precoce do espectro do autismo por muitos fatores **técnicos, científicos e sociais**, e a pesquisa apresentou ótimas contribuições para que este diagnóstico seja ainda mais preciso.

O projeto que o grupo realizou contribui muito para o esclarecimento do tema, principalmente para as pessoas que desconhecem sobre ele, gerando visibilidade positiva, além do incentivo necessário para novas pesquisas sobre TEA no meio acadêmico brasileiro. Sendo assim, se eu pudesse falar em nome da minha classe, psicólogos, quero agradecer pela responsabilidade com o tema e pela contribuição técnica-científica."

Psicóloga Juliana Arielo Tortorelli CRP 06/142415

Dúvidas?
