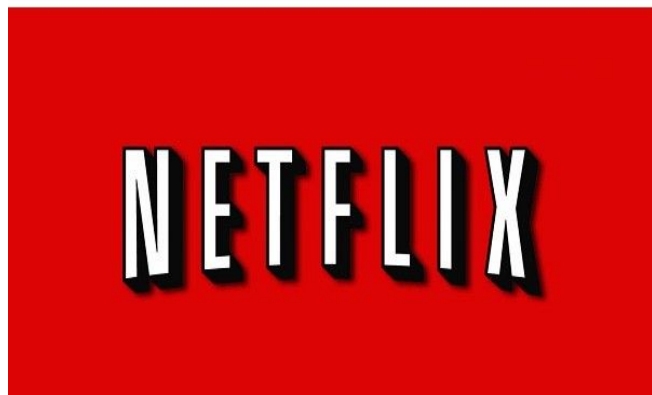


Recomendação de Conteúdo e QoE: Um Experimento Quantificando o Impacto da QoS nas Preferências por Conteúdos

Felipe Assis de Souza, Mateus Schulz Nogueira,
Daniel S. Menasche, João Ismael Pinheiro, Carla Delgado,
Pavlos Sermpezis, Thrasyvoulos Spyropoulos, Savvas Kastanakis

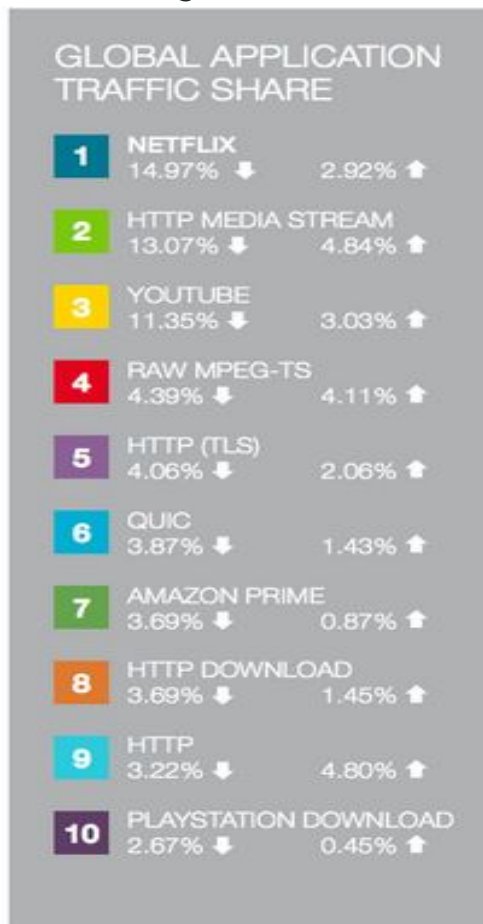
Parte 1: Introdução

Sistemas de Recomendação e Caches



Netflix atinge 130 milhões de assinaturas em Julho/2018[2]

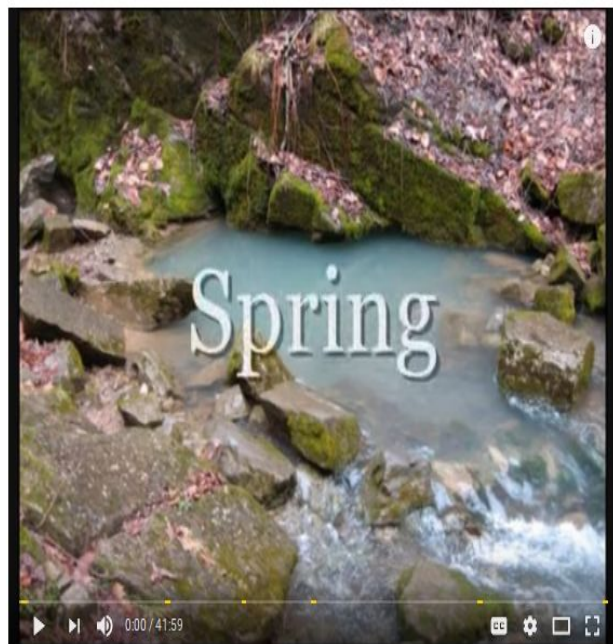
[2]<https://www.statista.com/chart/10311/netflix-subscriptions-usa-international/>



-Youtube tem 1.9 bilhões de usuários logados mensalmente.
-Mais de 1 bilhão de horas assistidas diariamente.[1]

[1]<https://www.youtube.com/yt/about/press/>

Sistemas de Recomendação



Four Seasons ~ Vivaldi

184,093,110 views

775K 40K SHARE SAVE ...

AnAmericanComposer

Up next

AUTOPLAY



Beethoven Silence
Tat Parina
38M views



Mix - Four Seasons ~ Vivaldi
YouTube



Music for reading - Chopin,
Beethoven, Mozart, Bach,
HALIDONMUSIC
22M views



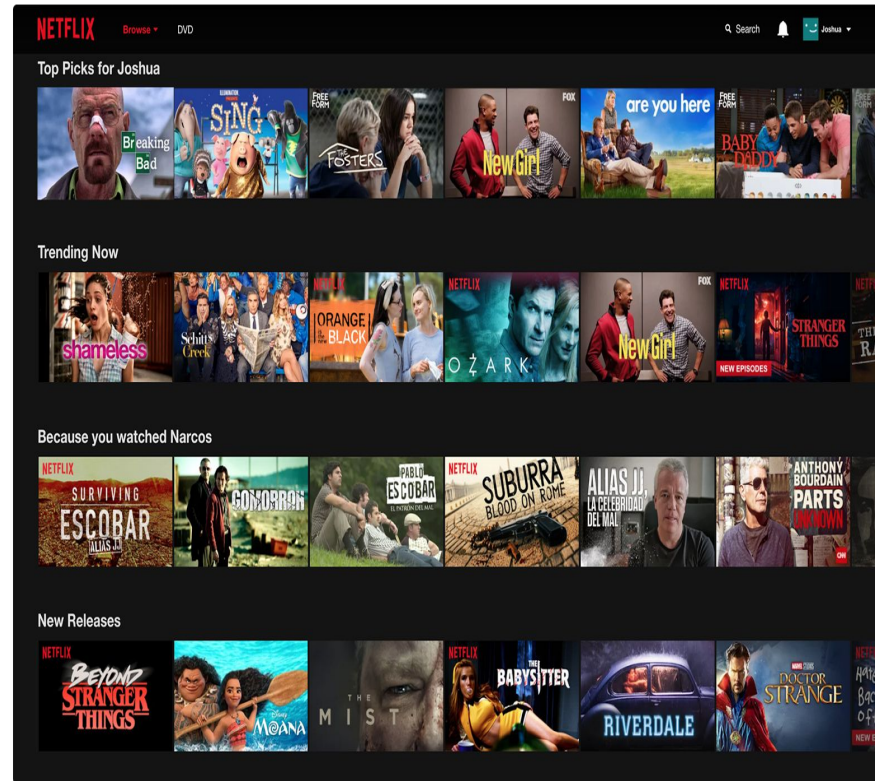
The Best of Classical Music -
Mozart, Beethoven, Bach,
Hawaii ASMR Nature Relaxation
815K views



Mozart Música Clássica para
Estudar e Memorizar | Músicas
Músicas Relaxantes
420K views



J.S. Bach: The Violin Concertos
Brilliant Classics
697K views



Motivação

Estima-se que o sistema de recomendação da Netflix influencia 80% da demanda.[3]

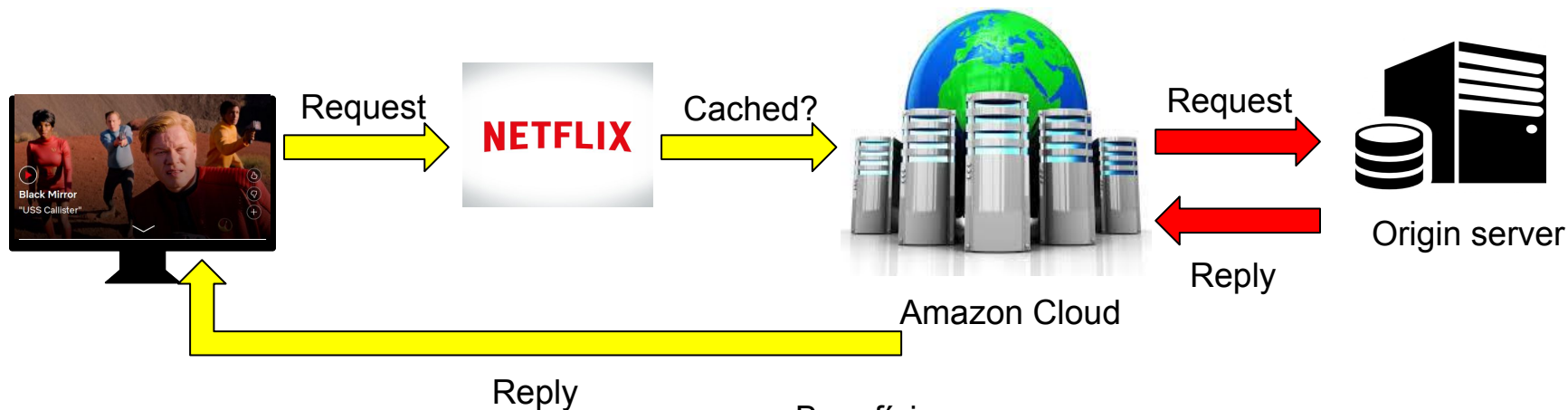


[3] Gomez-Uribe, Carlos A., and Neil Hunt. "The netflix recommender system: Algorithms, business value, and innovation." *ACM Transactions on Management Information Systems (TMIS)* 6.4 (2016): 13.

O que considerar num sistema de recomendação?

1. ideia 1
2. ideia 2
3.
4. ...
- 5.

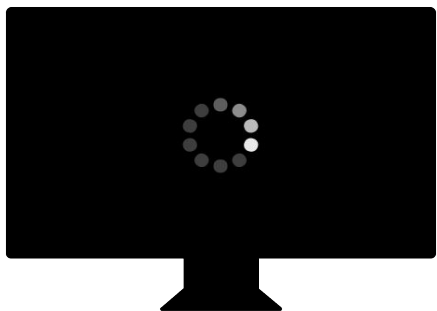
Sistemas de Caches



Benefícios:

- Consumo de banda reduzido
- Prevenção de atrasos de transmissão

O Experimento



VS



Qual deles um usuário iria preferir?

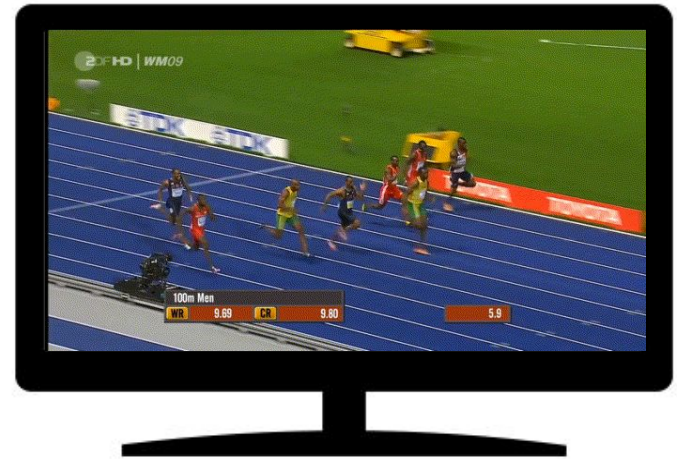
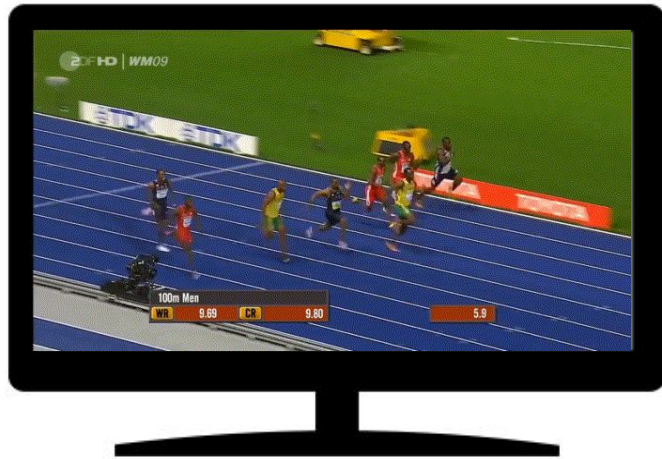
Questões a serem respondidas

- Quais são os critérios utilizados por um usuário ao escolher um vídeo?
- É possível quantificar o quanto um usuário é sensível a falhas na qualidade da transmissão?
- É possível quantificar o interesse do usuário por algum conteúdo?
- O que é mais importante, o conteúdo do vídeo ou a qualidade da transmissão?



Metodologia

- Experimento dividido em 3 etapas:
 - a. O usuário responde um rápido questionário no qual ordena 5 categorias gerais de conteúdo em ordem de preferência (esportes, comédia, música, documentários e animais).
 - b. O usuário assiste a um par de vídeos cujos conteúdos são iguais, diferindo apenas na qualidade.
 - c. O usuário assiste a outros 3 pares de vídeos, sorteados aleatoriamente dentre um conjunto de 7 pares.
 - Cada par é constituído de modo que um dos vídeos seja mais interessante do que o outro. O vídeo mais interessante possui uma qualidade inferior ao vídeo menos interessante.



Vídeos com o mesmo conteúdo mas com qualidade diferente.



Vídeos com conteúdo e qualidade distintas.

Dados Coletados

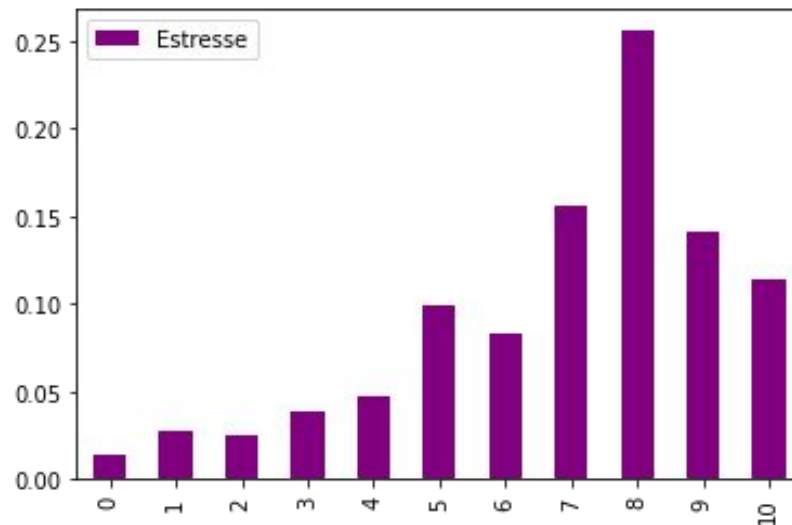
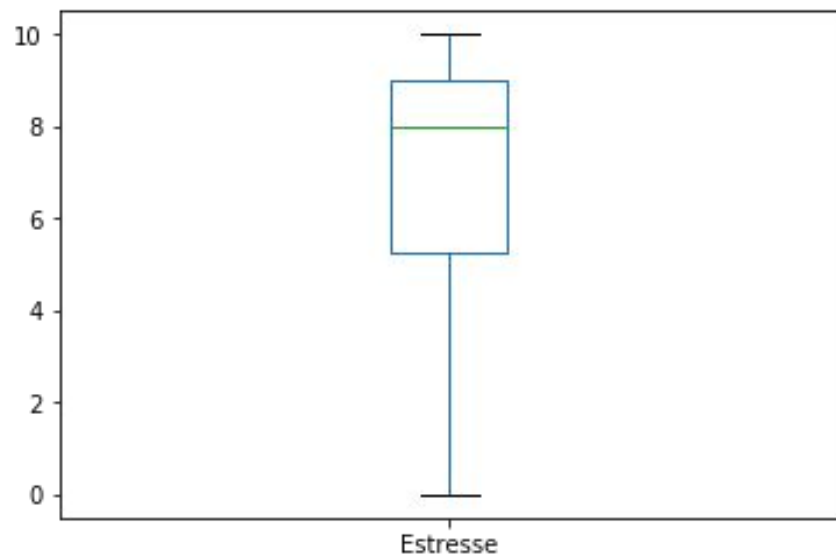
- **Estresse:** o quanto o usuário se sentiu afetado pelas falhas na qualidade de serviço. *[numérico de 0 a 10]*
- **Interesse VideoLowQoS:** Interesse pelo conteúdo do vídeo de qualidade baixa. *[numérico de 0 a 5]*
- **Interesse VideoHighQoS:** Interesse pelo conteúdo do vídeo de qualidade alta. *[numérico de 0 a 5]*
- **Incômodo:** indica o quanto o usuário se sentiu incomodado com as falhas. *[numérico de 0 a 5]*
- **Vídeo Preferido:** Qual dos vídeos de uma sessão foi preferido pelo usuário. *[variável categórica]*
- **Justificativa:** informação textual.

Ferramentas utilizadas



Parte 2: Processamento dos Dados

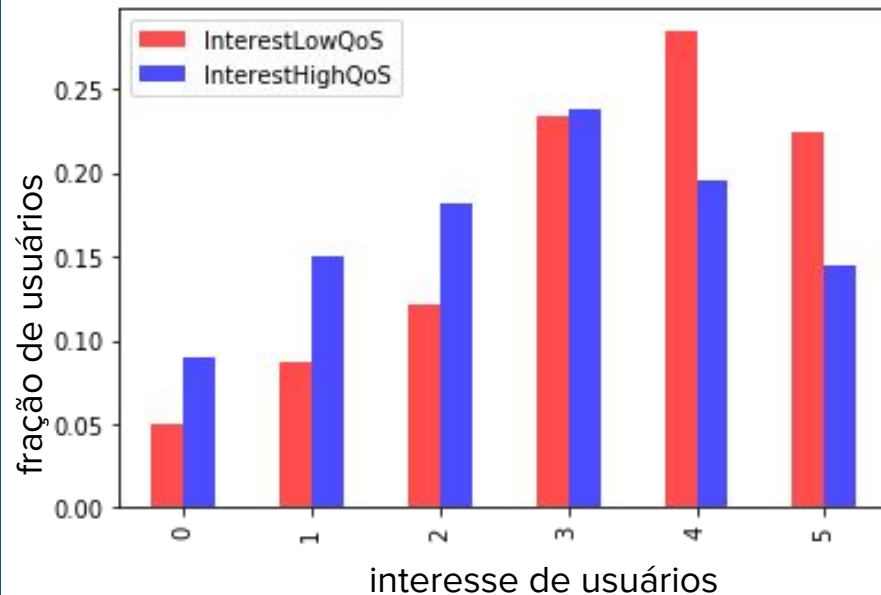
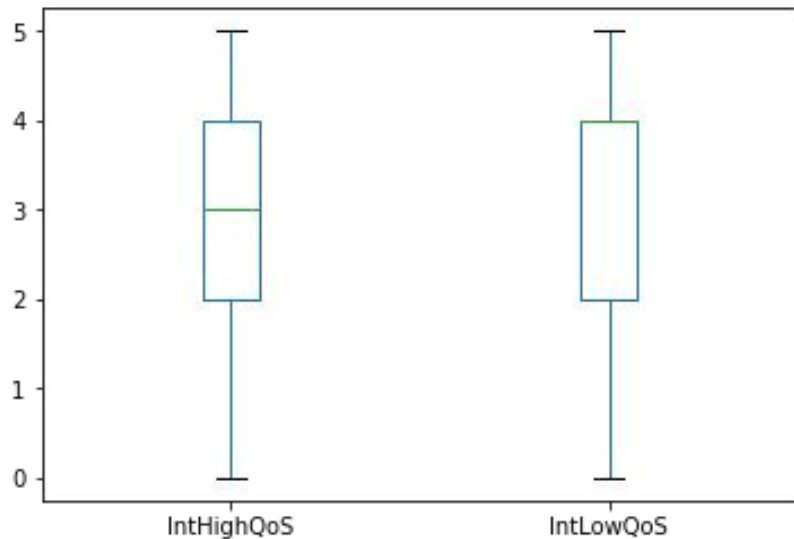
Estresse



Média	Desvio Padrão
6.922819	2.369949

As falhas de QoS não passaram despercebidas.

Interesses



	Média	Desvio Padrão
IntLowQoS	3.285235	1.422214
IntHighQoS	2.733781	1.513892

Existe um *tradeoff* entre qualidade de transmissão e interesse pelos vídeos.

Preferências dos usuários por categoria

Description	Prefer High QoS (total)	Interest Low Qos	Interest High QoS	Annoy- ance
All	561 (1002)	3.31 ± 1.42	2.77 ± 1.51	3.56 ± 1.33
1) Soccer	75 (132)	3.27 ± 1.64	2.50 ± 1.66	3.83 ± 1.19
2) Comedy	78 (118)	3.26 ± 1.34	2.97 ± 1.50	3.52 ± 1.42
3) Pets	64 (135)	2.93 ± 1.52	2.13 ± 1.51	3.32 ± 1.42
4) Sports	84 (158)	3.16 ± 1.35	2.49 ± 1.35	3.53 ± 1.24
5) Animals	95 (135)	3.13 ± 1.38	3.08 ± 1.49	3.58 ± 1.39
6) Science	55 (141)	3.94 ± 1.25	3.05 ± 1.43	3.33 ± 1.33
7) Music	98 (163)	3.56 ± 1.22	3.18 ± 1.40	3.80 ± 1.30

Preferências dos usuários por categoria

A maioria dos usuários preferiu o vídeo de alta QoS

	Prefer High QoS (total)	Interest Low Qos	Interest High QoS	Annoy-ance
	561 (1002)	3.31 ± 1.42	2.77 ± 1.51	3.56 ± 1.33
1) Soccer	75 (132)	3.27 ± 1.64	2.50 ± 1.66	3.83 ± 1.19
2) Comedy	78 (118)	3.26 ± 1.34	2.97 ± 1.50	3.52 ± 1.42
3) Pets	64 (135)	2.93 ± 1.52	2.13 ± 1.51	3.32 ± 1.42
4) Sports	84 (158)	3.16 ± 1.35	2.49 ± 1.35	3.53 ± 1.24
5) Animals	95 (135)	3.13 ± 1.38	3.08 ± 1.49	3.58 ± 1.39
6) Science	55 (141)	3.94 ± 1.25	3.05 ± 1.43	3.33 ± 1.33
7) Music	98 (163)	3.56 ± 1.22	3.18 ± 1.40	3.80 ± 1.30

Preferências dos usuários por categoria

A maioria dos usuários preferiu o vídeo de alta QoS

	Prefer High QoS (total)	Interest Low Qos	Interest High QoS	Annoy-ance
	561 (1002)	3.31 ± 1.42	2.77 ± 1.51	3.56 ± 1.33
1) Soccer	75 (132)	3.27 ± 1.64	2.50 ± 1.66	3.83 ± 1.19
2) Comedy	78 (118)	3.26 ± 1.34	2.97 ± 1.50	3.52 ± 1.42
3) Pets	64 (135)	2.93 ± 1.52	2.13 ± 1.51	3.32 ± 1.42
4) Sports	84 (158)	3.16 ± 1.35	2.49 ± 1.35	3.53 ± 1.24
5) Animals	95 (135)	3.13 ± 1.38	3.08 ± 1.49	3.58 ± 1.39
6) Science	55 (141)	3.94 ± 1.25	3.05 ± 1.43	3.33 ± 1.33
7) Music	98 (163)	3.56 ± 1.22	3.18 ± 1.40	3.80 ± 1.30

Embora o interesse maior fosse pelo vídeo de Low QoS

Preferências dos usuários por categoria

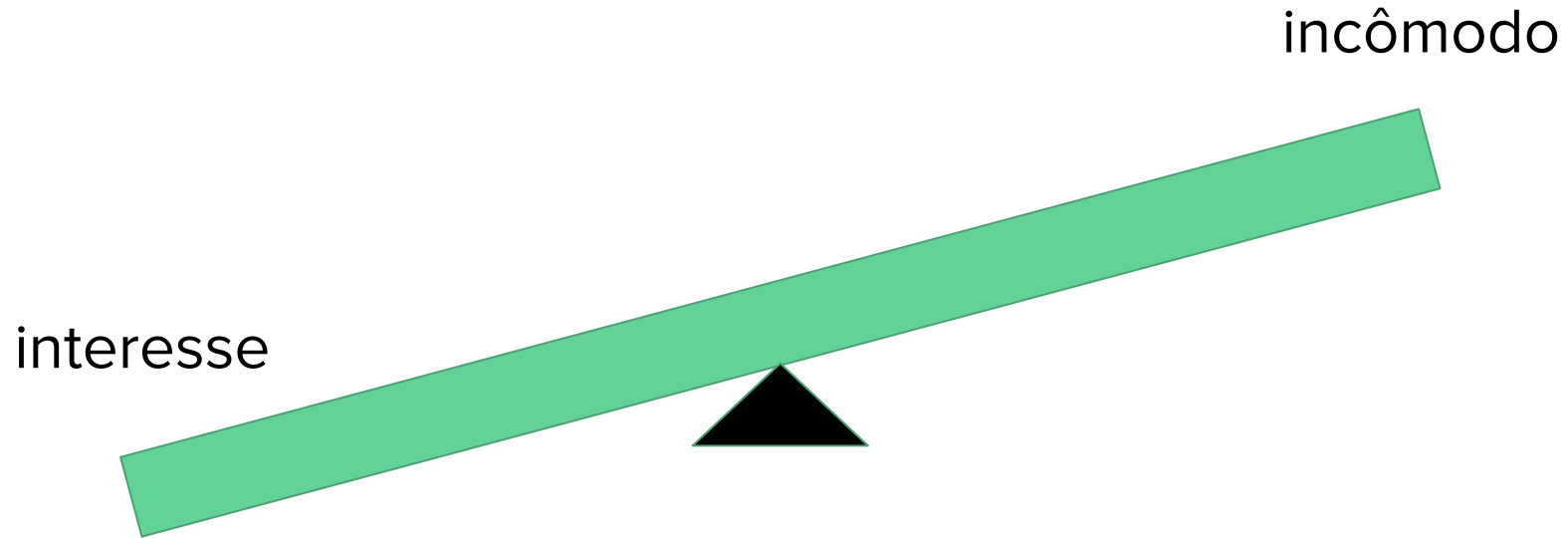


A maioria dos usuários preferiu o vídeo de alta QoS

	Prefer High QoS (total)	Interest Low Qos	Interest High QoS	Annoy-ance
	561 (1002)	3.31 ± 1.42	2.77 ± 1.51	3.56 ± 1.33
1) Soccer	75 (132)	3.27 ± 1.64	2.50 ± 1.66	3.83 ± 1.19
2) Comedy	78 (118)	3.26 ± 1.34	2.97 ± 1.50	3.52 ± 1.42
3) Pets	64 (135)	2.93 ± 1.52	2.13 ± 1.51	3.32 ± 1.42
4) Sports	84 (158)	3.16 ± 1.35	2.49 ± 1.35	3.53 ± 1.24
5) Animals	95 (135)	3.13 ± 1.38	3.08 ± 1.49	3.58 ± 1.39
6) Science	55 (141)	3.94 ± 1.25	3.05 ± 1.43	3.33 ± 1.33
7) Music	98 (163)	3.56 ± 1.22	3.18 ± 1.40	3.80 ± 1.30

Embora o interesse maior fosse pelo vídeo de Low QoS

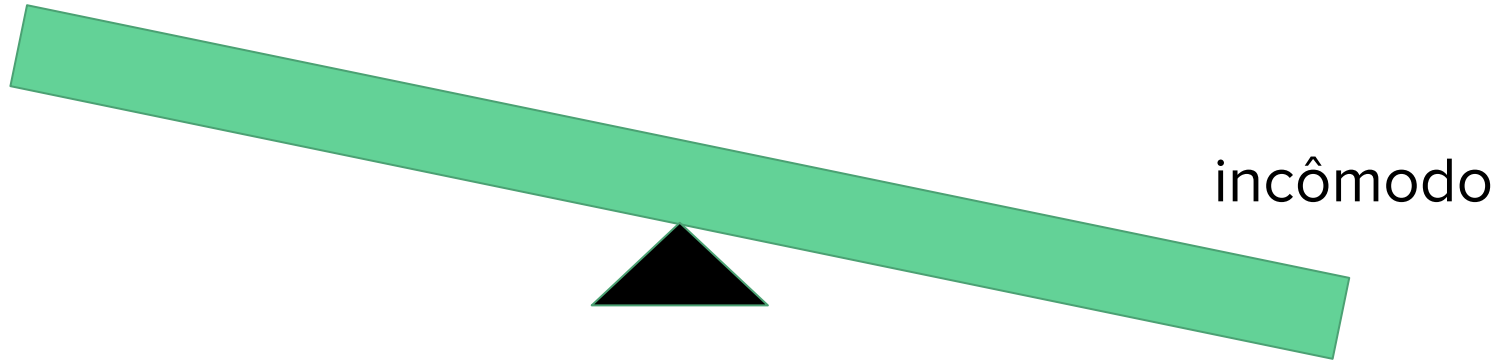
Tradeoff: Interesse versus Incômodo



Resultado: usuário prefere vídeo de baixa QoS (Low QoS)

Tradeoff: Interesse versus Incômodo

interesse

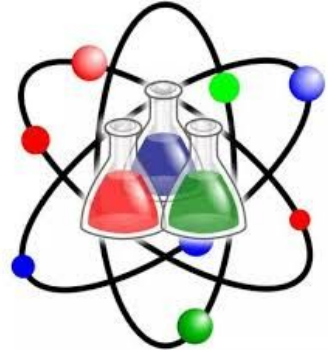
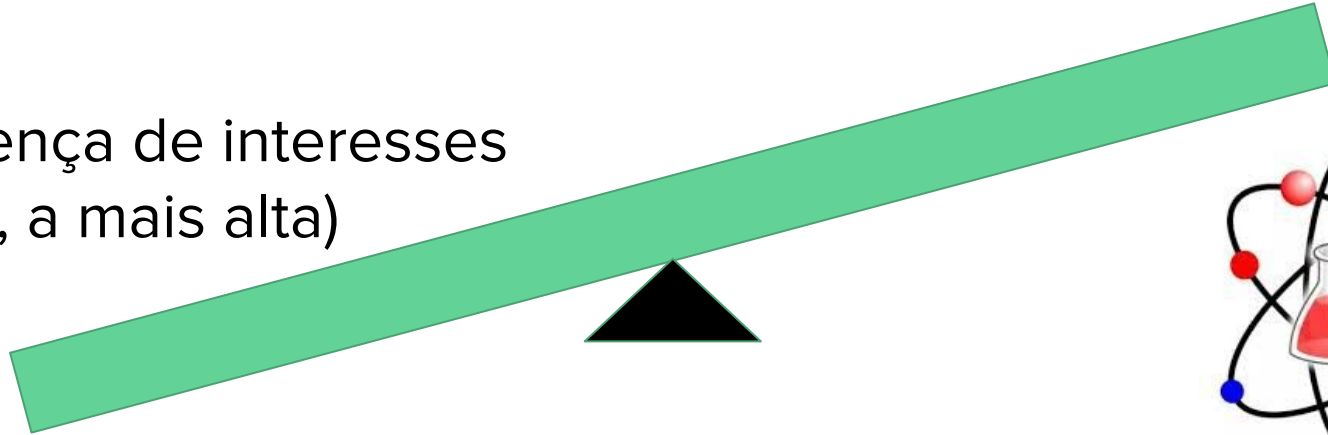


Resultado: usuário prefere vídeo de alta QoS (High QoS)

Ciência: Interesse Alto Impacto

incômodo (3.33, o
menor observado)

diferença de interesses
(0.90, a mais alta)



Resultado: usuário prefere vídeo de baixa QoS (Low QoS)

Música: Incômodo Alto Impacto

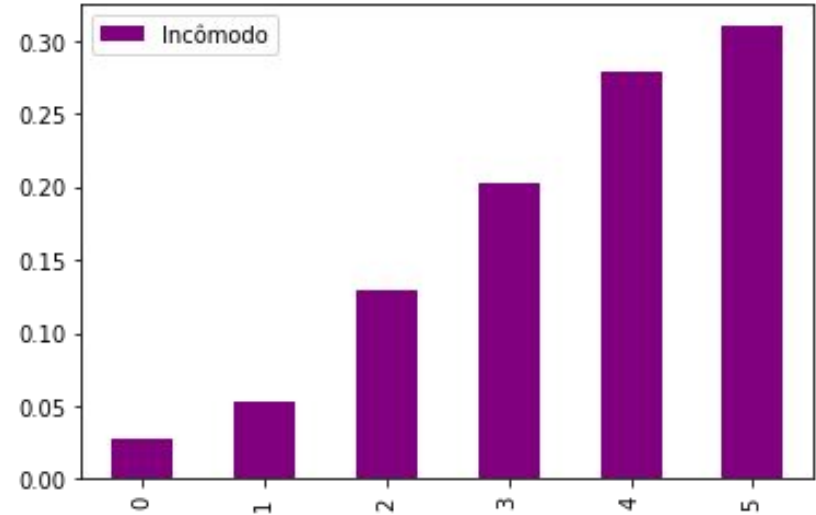
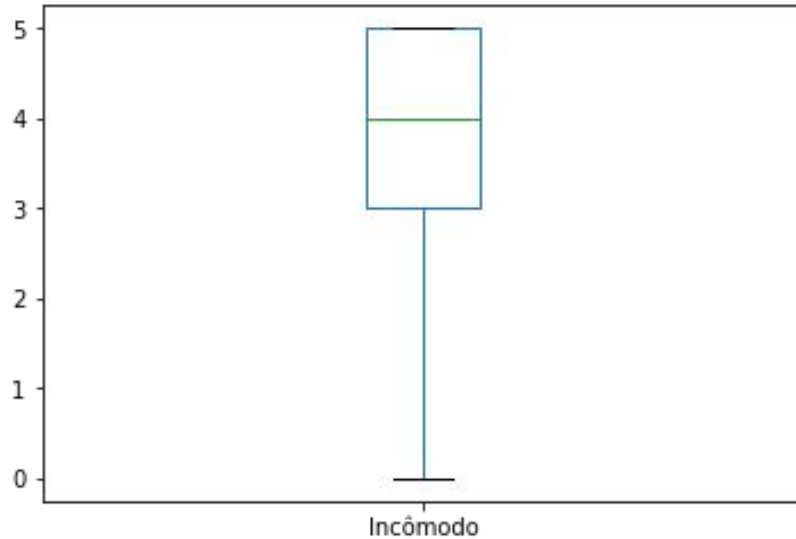
diferença de interesses
(0.38)



incômodo (3.8)

Resultado: usuário prefere vídeo de alta QoS (High QoS)

Incômodo



Média	Desvio Padrão
3.585011	1.326379

Os usuários se incomodam bastante com distúrbios de qualidade

Estatísticas por sessão [1/4]

Sessão 1 [139]	IntLowQoS	IntHighQoS	Incômodo
Média	3.086331	2.42446	3.755396
Desvio padrão	1.617307	1.592457	1.184695

LowQoS	43%
HighQoS	57%

esporte



Falhas para esportes pode incomodar muito mais.

Sessão 2 [109]	IntLowQoS	IntHighQoS	Incômodo
Média	3.238532	2.899083	3.559633
Desvio padrão	1.353316	1.403905	1.545273

LowQoS	31%
HighQoS	69%

comédia



68% preferiram o vídeo com maior qualidade.

Estatísticas por sessão [2/4]

Sessão 3 [122]	IntLowQoS	IntHighQoS	Incômodo
Média	2.868852	2.098361	3.303279
Desvio padrão	1.520933	1.501569	1.430999

LowQoS	52%
HighQoS	48%

pets



Menor interesse para a categoria de animais.

Sessão 4 [141]	IntLowQoS	IntHighQoS	Incômodo
Média	3.141844	2.489362	3.588652
Desvio padrão	1.376235	1.381599	1.248261

LowQoS	47%
HighQoS	53%

esporte
radical



Equilíbrio na preferência de vídeos apesar da diferença expressiva de interesses.

Estatísticas por sessão [3/4]

Sessão 5 [118]	IntLowQoS	IntHighQoS	Incômodo
Média	3.076271	2.983051	3.694915
Desvio padrão	1.378447	1.461664	1.349317

LowQoS	28%
HighQoS	72%

animais



Maior proporção de usuários que preferiram qualidade alta (72%).

Sessão 6 [124]	IntLowQoS	IntHighQoS	Incômodo
Média	3.935484	3.056452	3.306452
Desvio padrão	1.228006	1.438735	1.301515

LowQoS	60%
HighQoS	40%

ciência



Muitos preferiram menor qualidade.
Segunda menor média de incômodo.

Estatísticas por sessão [4/4]

Sessão 7 [141]	IntLowQoS	IntHighQoS	Incômodo
Média	3.624113	3.212766	3.829787
Desvio padrão	1.168265	1.382588	1.319943

LowQoS	38%
HighQoS	62%

música

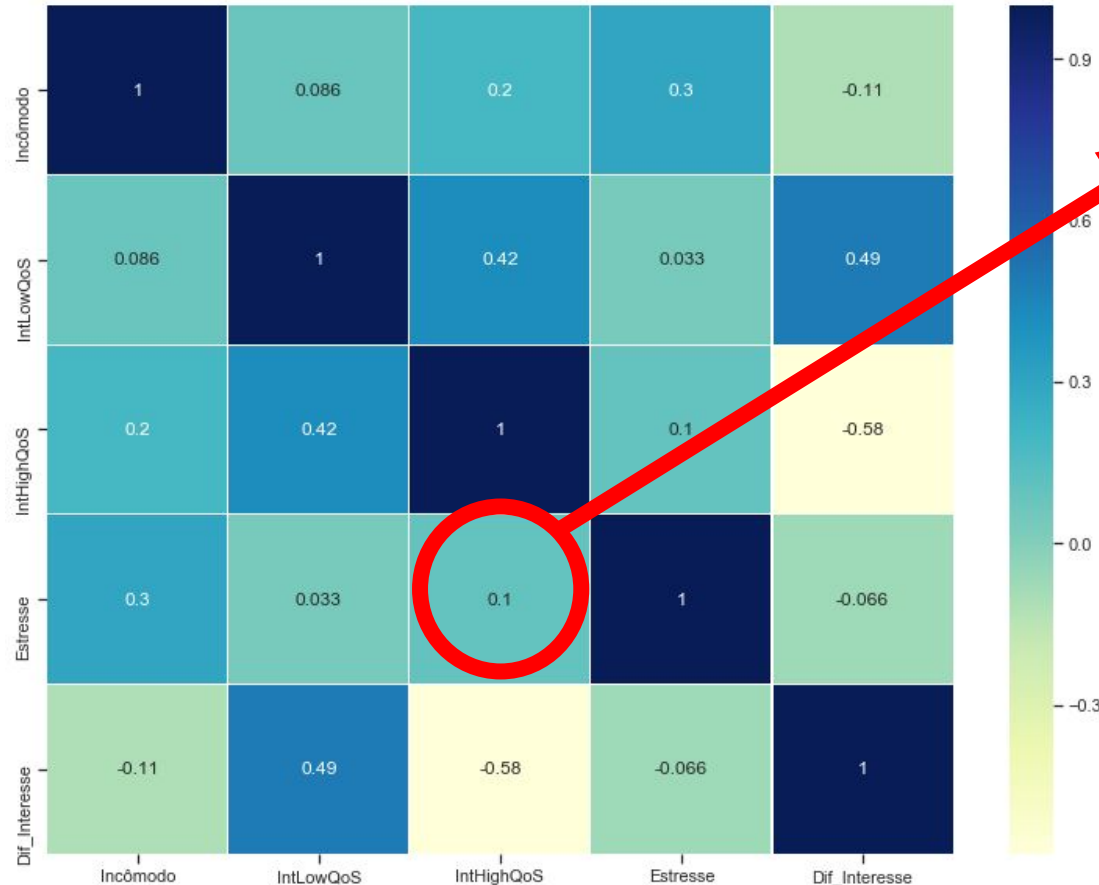


A média do incômodo para a categoria música foi a mais alta.

Em todas as sessões o vídeo com pior qualidade foi considerado mais interessante.

Nas duas sessões em que a maioria preferiu o vídeo de pior qualidade foram observadas as menores médias de incômodo

Matriz de Correlações

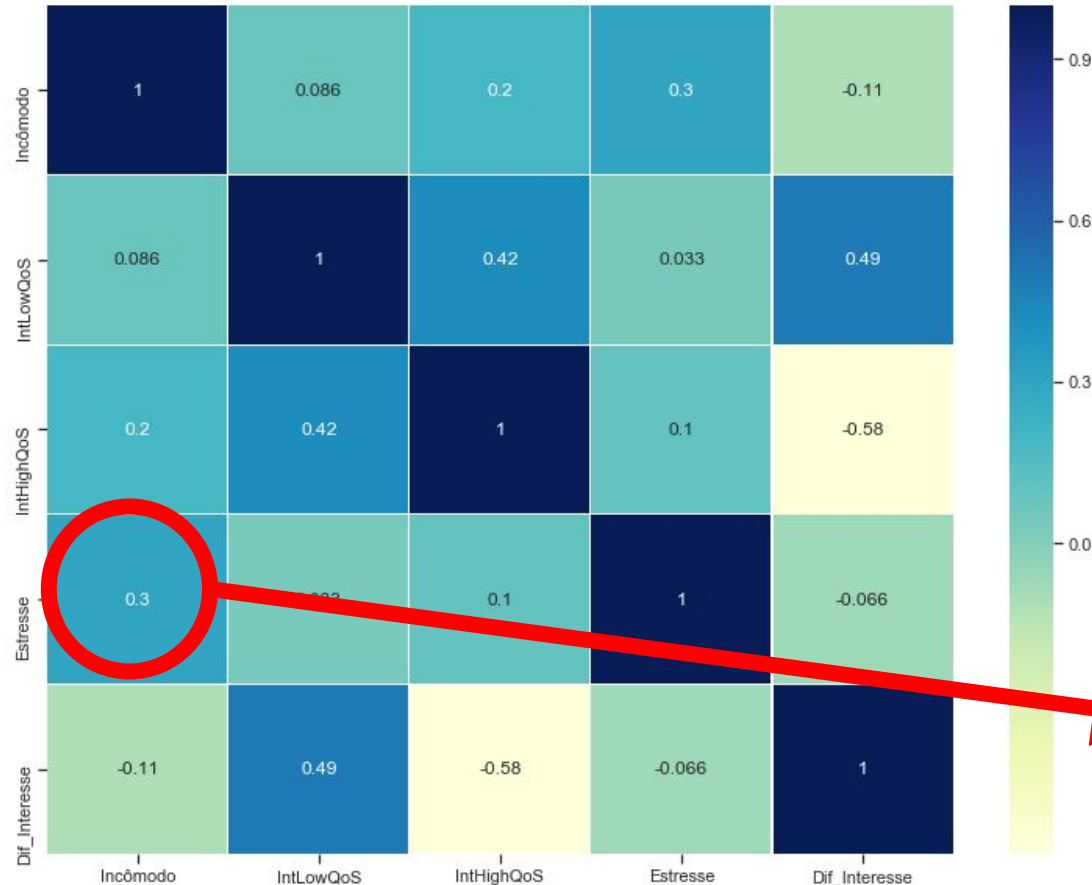


Correlação baixa entre [estresse, incômodo] e [IntHighQoS, IntLowQoS].



Usuários entenderam a distinção entre interesse pelo vídeo e incômodo com a qualidade

Matriz de Correlações



Correlação baixa entre [estresse, incômodo] e [IntHighQoS, IntLowQoS].



Usuários entenderam a distinção entre interesse pelo vídeo e incômodo com a qualidade

Entre estresse e incômodo há alta correlação

Wordcloud



Nuvem de palavras mais frequentes da classe LowQoS



Nuvem de palavras mais frequentes da classe HighQoS

Wordcloud



(a) Prefer high QoS video

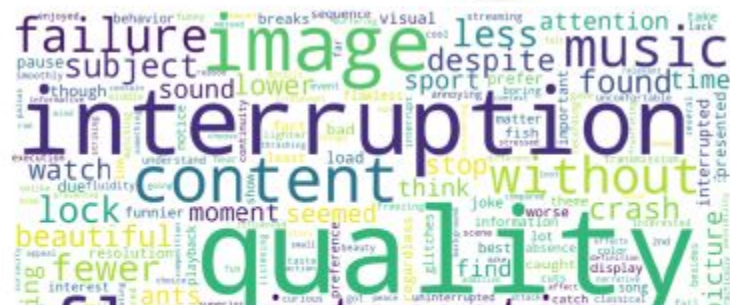


(b) Prefer low QoS video



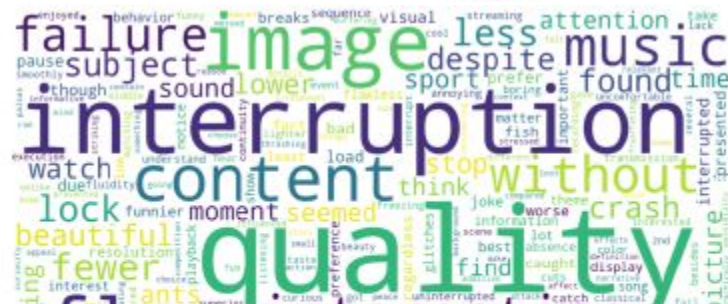
(b) Prefer low QoS video

Wordcloud



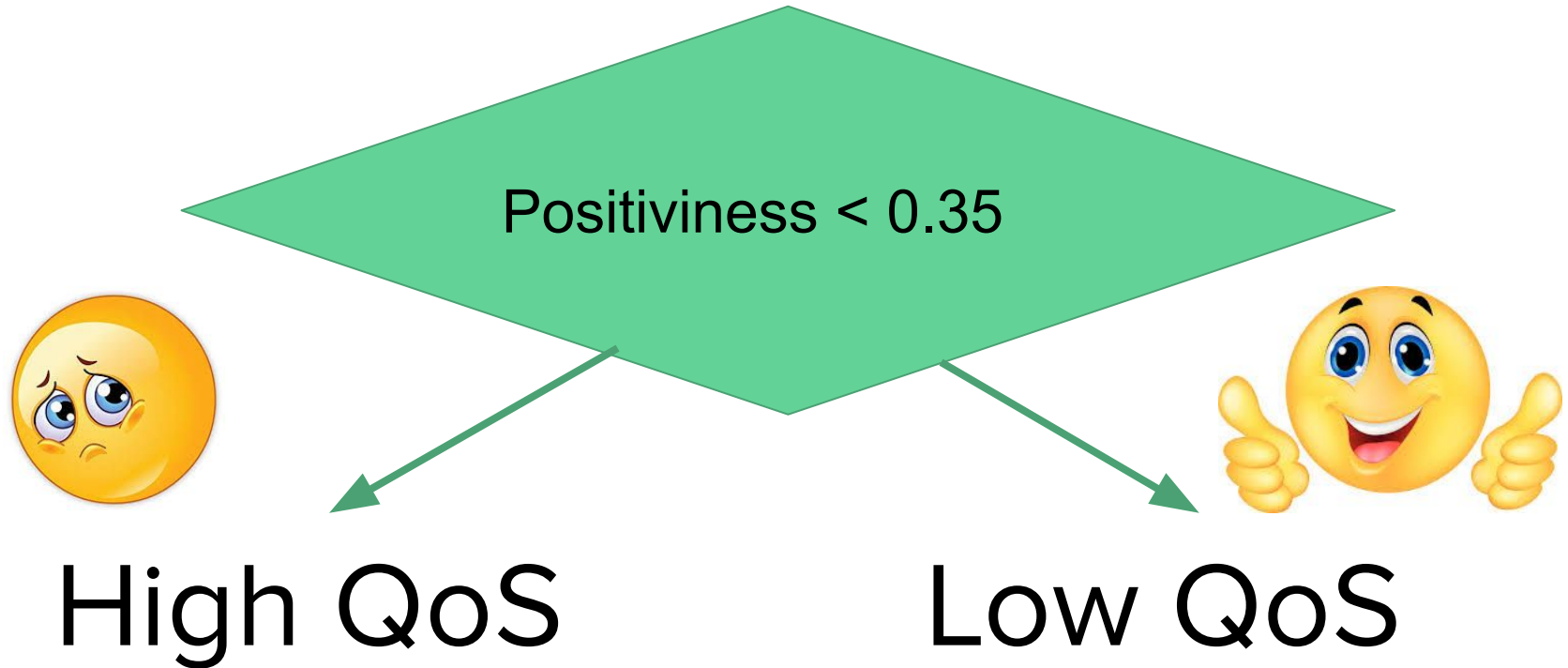
Can we use sentiment analysis to automatically classify users?

Wordcloud



Can we use sentiment analysis to automatically classify users?

Sentiment analysis: accuracy of 62%

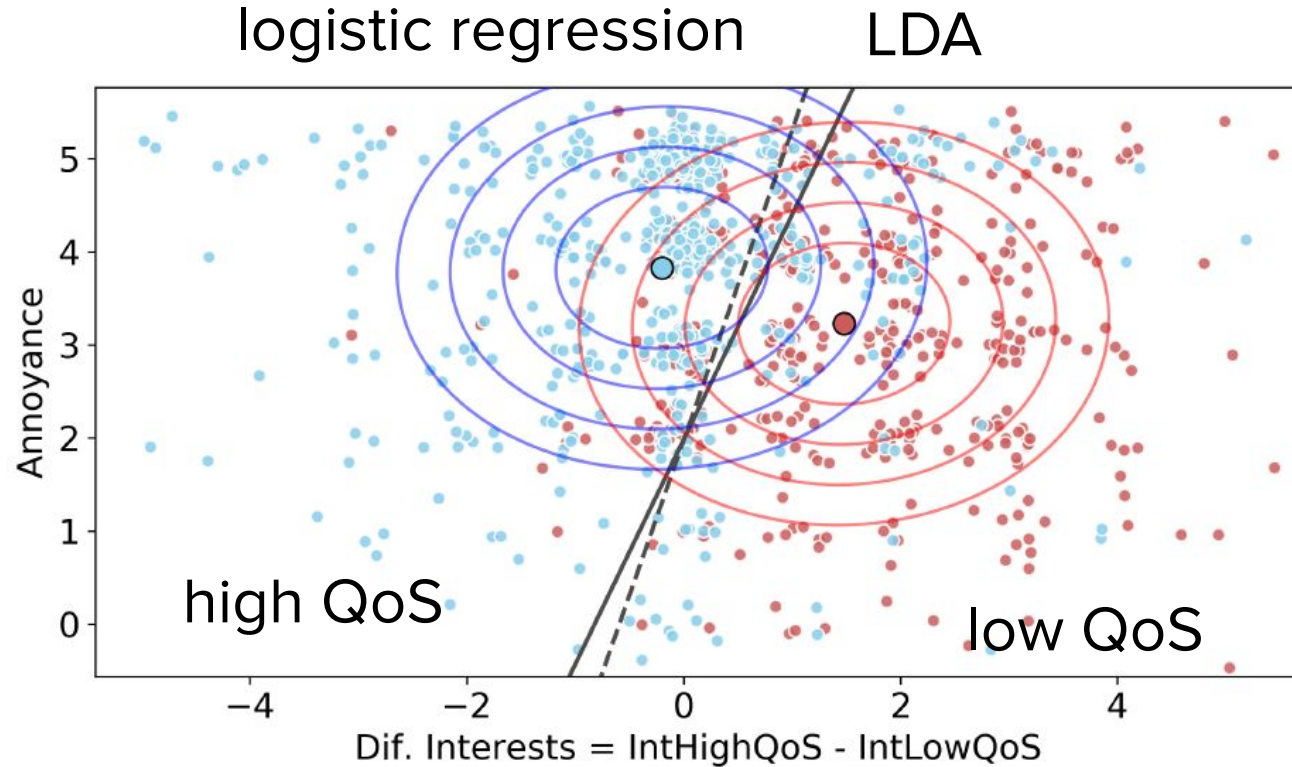


Natural Language Processing

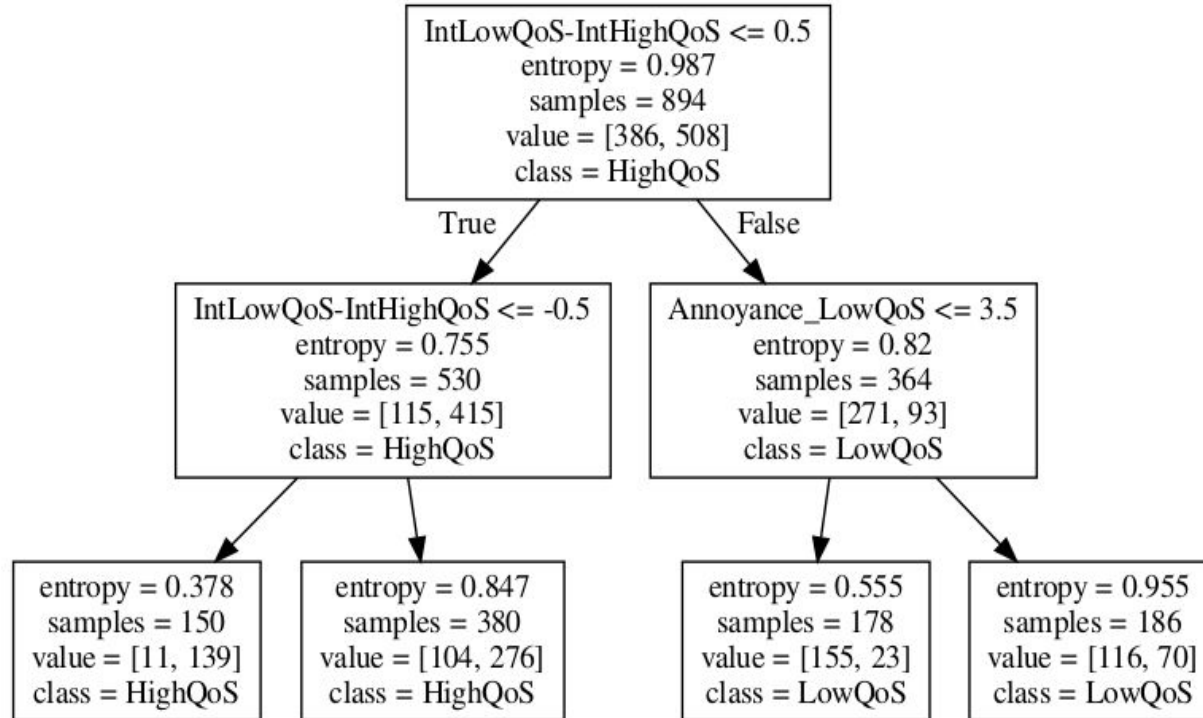
- tentar descobrir se usuário fez escolha por
 - rejeição (detestou o outro)
 - vontade (porque gostou muito daquela)
- melhorar nuvens de pontos
 - criar classes representativas, por exemplo, pegando extremos
 - tirar palavras irrelevantes
- fazer análise de sentimento e aplicar outros algoritmos padrões de NLP nos dados
- etiquetar alguns (todos?!) manualmente, criando features extras, e ver se consegue algum nível de classificação automática

Parte 3: Modelos e Resultados

Visualização dos dados



Árvore de Decisão [1]

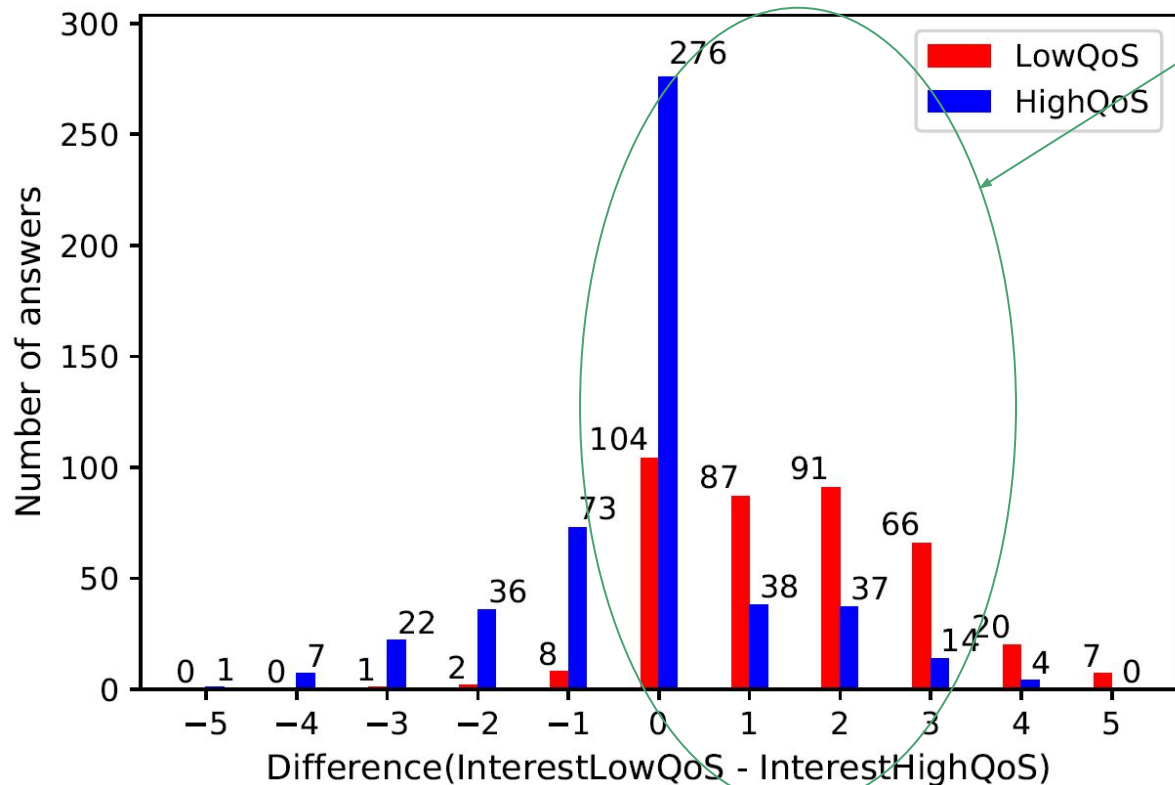


min_samples_split	38
min_samples_leaf	11
max_depth	3

O interesse nos vídeos se mostrou importante

Preferência x Diferença de Interesse

focus attention here

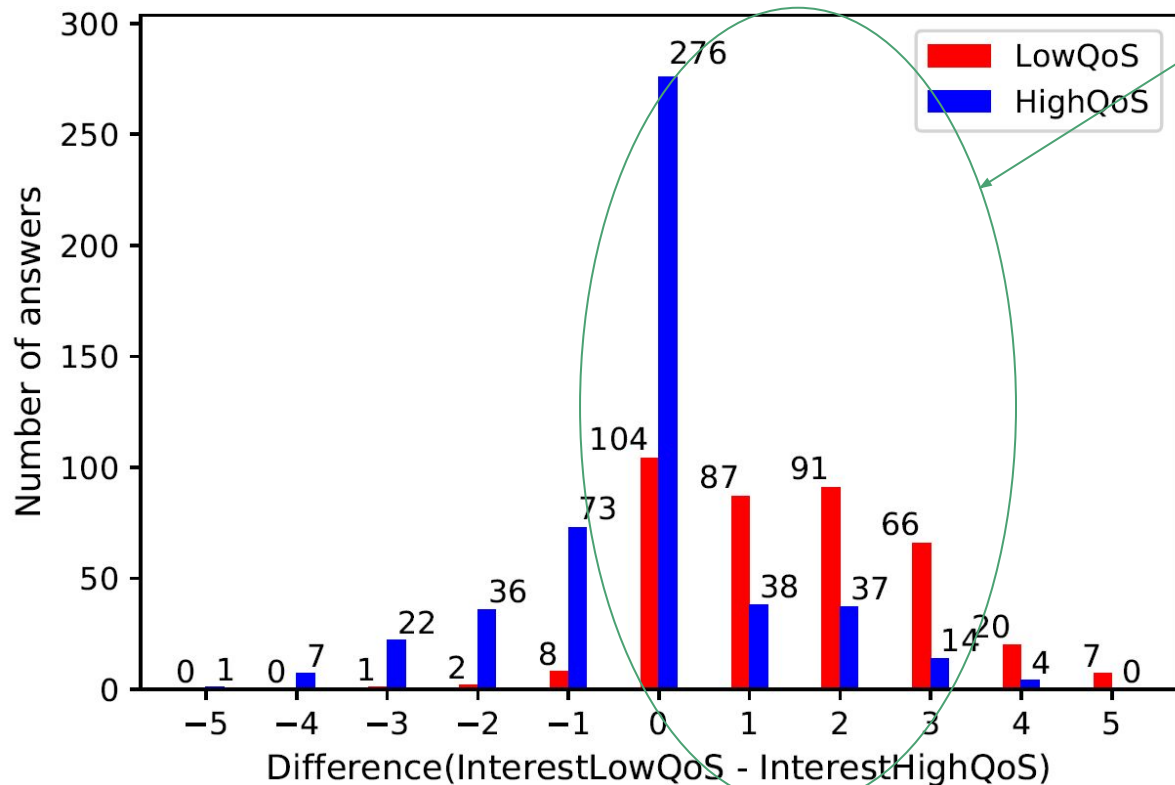


QoS é muito importante quando a diferença de interesse pelos vídeos é nula.

Interesse é muito importante mas não é absoluto.

Preferência x Diferença de Interesse

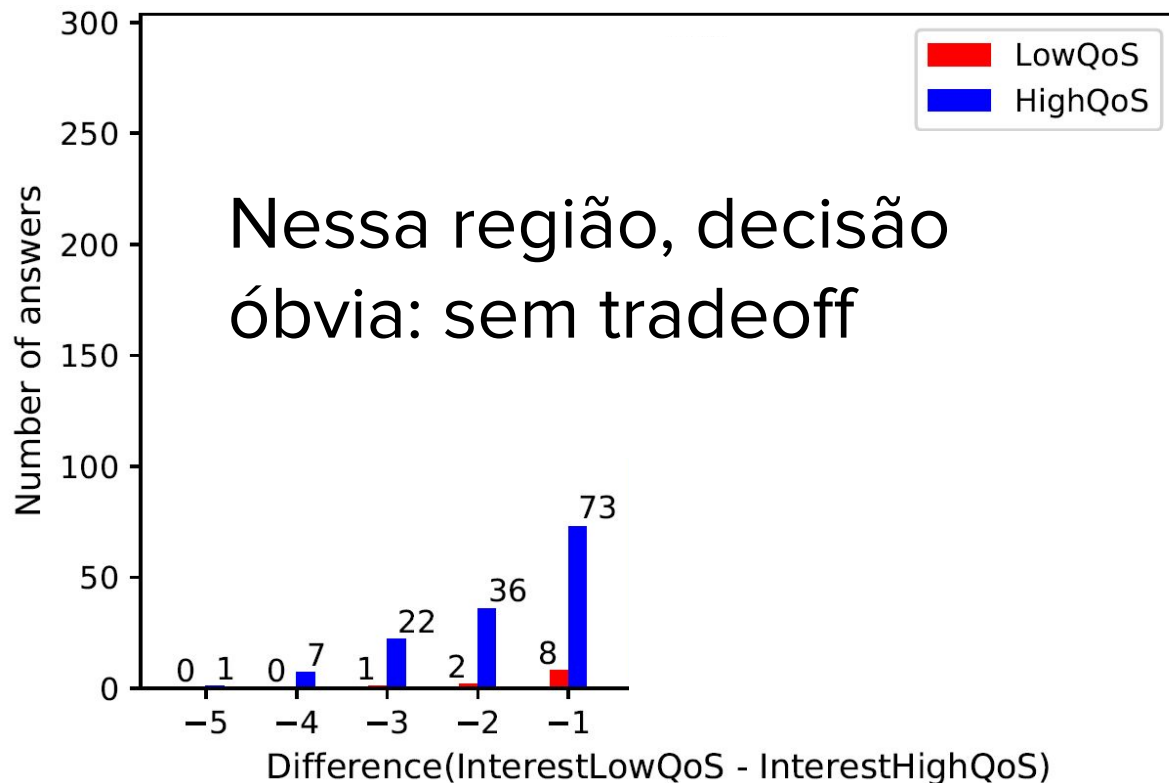
focus attention here



QoS é muito importante quando a diferença de interesse pelos vídeos é nula.

Interesse é muito importante mas não é absoluto.

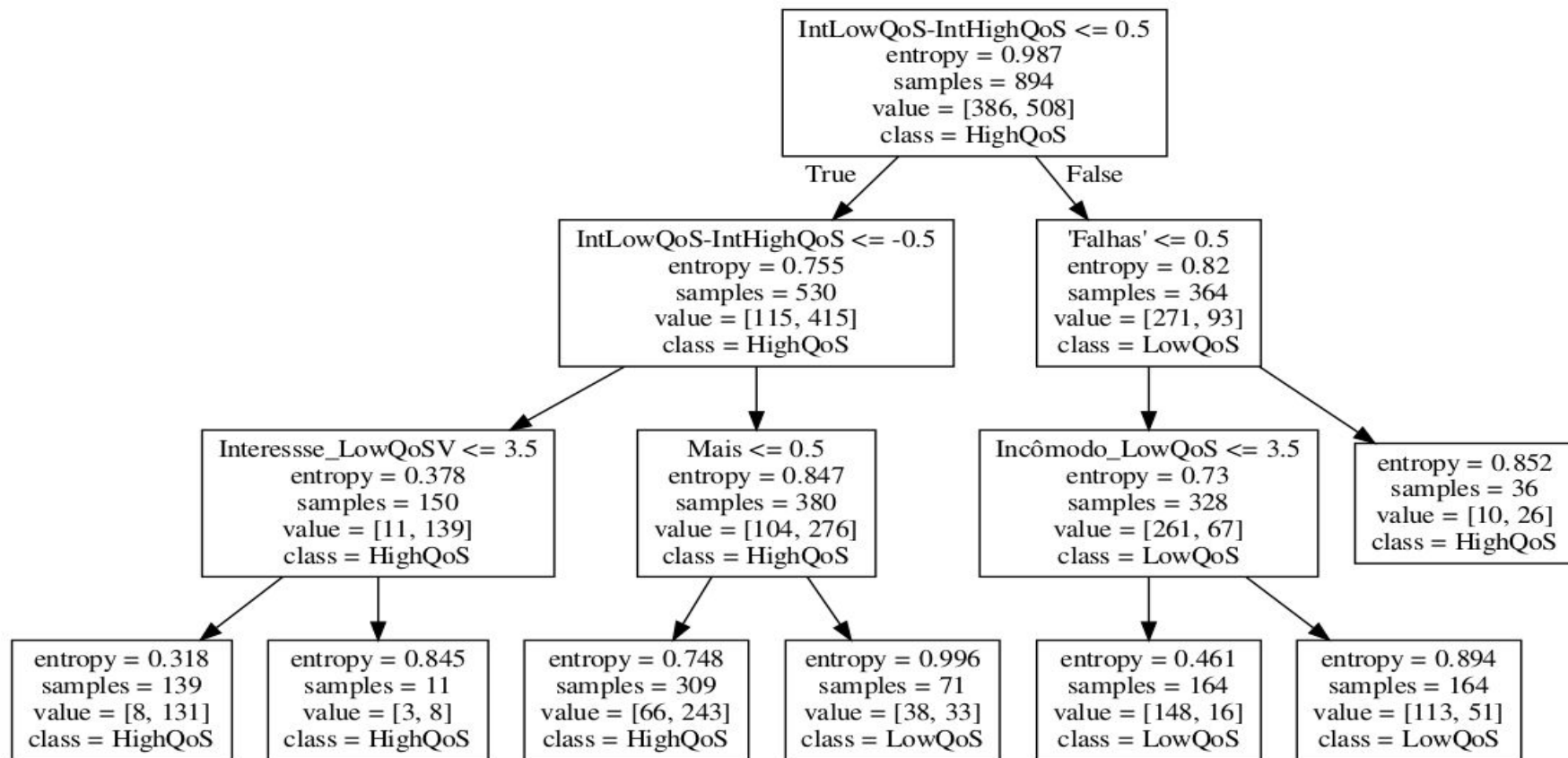
Preferência x Diferença de Interesse



Challenges

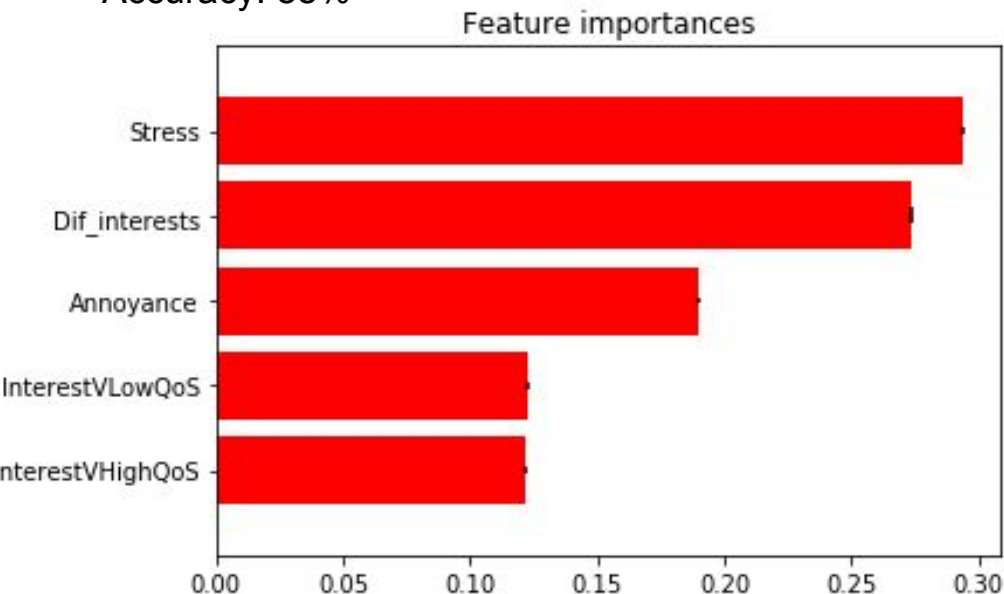
- unbalanced classes
 - use weights to fix
 - however, "accuracy" decreased
 - report multiple metrics of "accuracy"
 - use bootstrapping techniques
 - use different cost matrices, to account for false positives and false negatives
- check table here
 - https://en.wikipedia.org/wiki/Precision_and_recall
- interesting observation
 - the root of the new decision tree is now "annoyance" (high QoS video may be preferable if "annoyance is small") - only if "annoyance" is very high users prefer the high QoS

Árvore de Decisão com análise textual



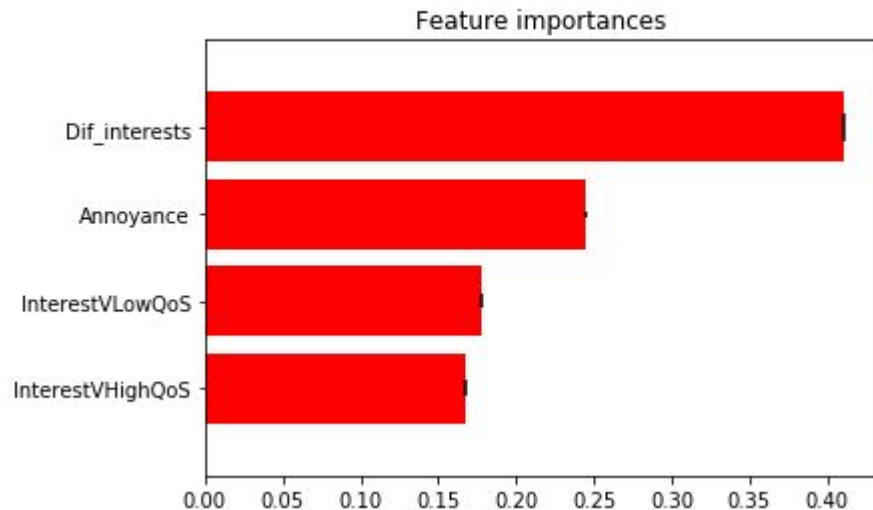
Extra Trees Classifier (check also xgboost and gradient boosting)

Accuracy: 88%



Without Stress(8% of accuracy decrease)

0.7986577181208053



Conclusão

1. Propusemos um experimento para quantificar a sensibilidade dos usuários ao conteúdo de vídeos e a qualidade de transmissão (QoS).
2. Obtivemos mais de **1,000** amostras de usuários reais sobre seus sentimentos com relação aos vídeos.
3. Propusemos um sistema de recomendação que leva em conta a qualidade da transmissão dos vídeos (QoS) e a natureza do conteúdo recomendado (QoR), conjuntamente impactando a qualidade de experiência (QoE).

Link para o experimento: www.experimentoqos.com

mateussnogs@gmail.com

felipe.s@ufrj.br

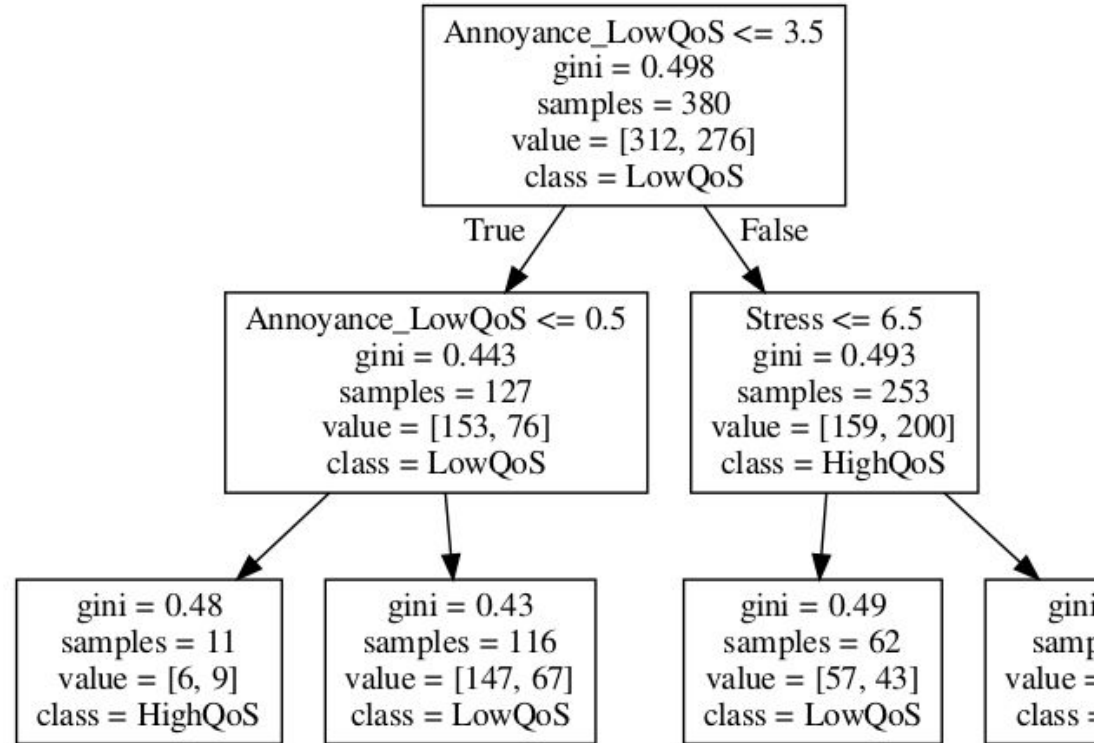
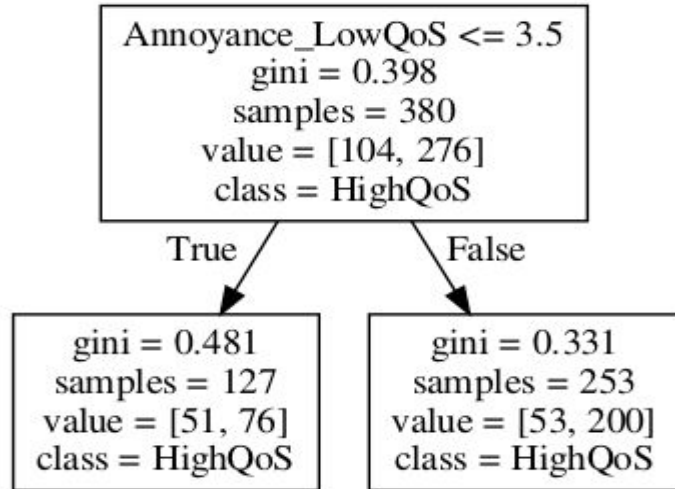
Fim

Obrigado!

Tree Dif = 0

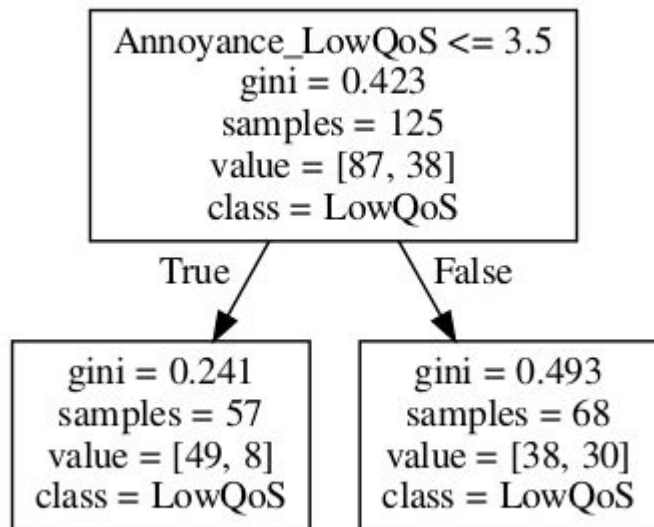
Balanced

Unbalanced(No cotnribution)

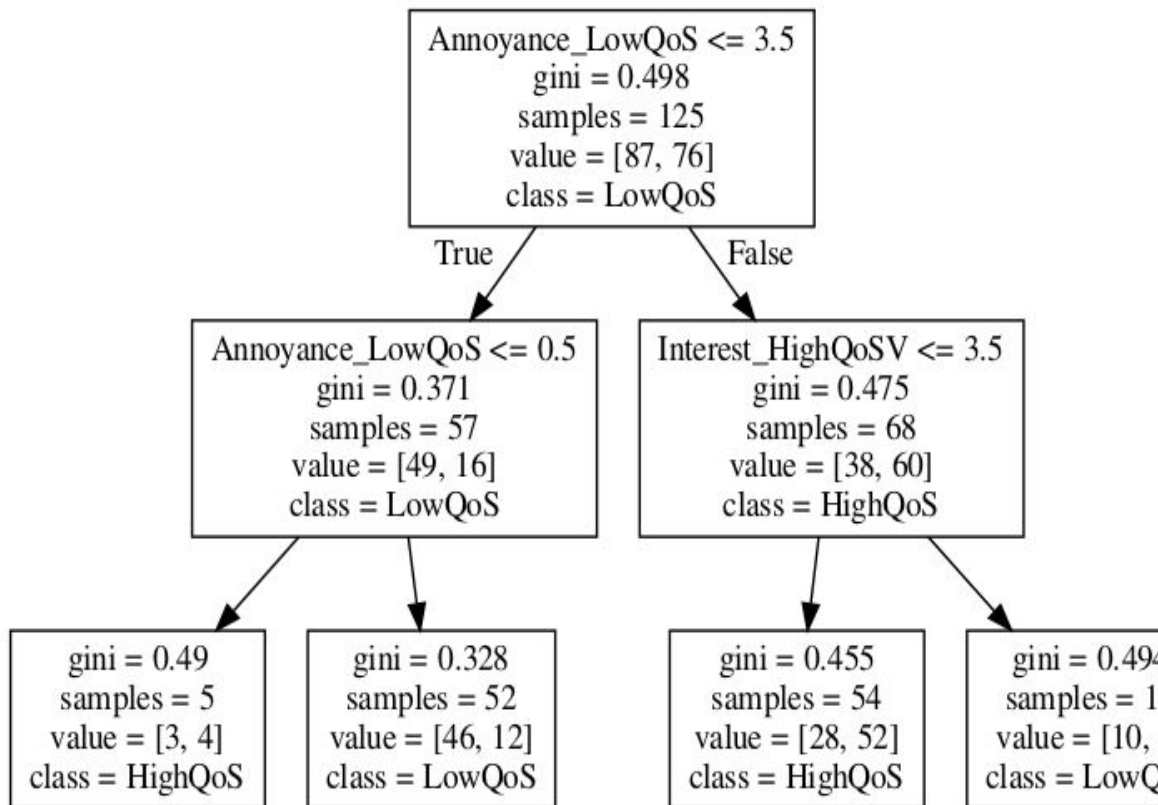


Tree Dif = 1

Unbalanced(no contribution)

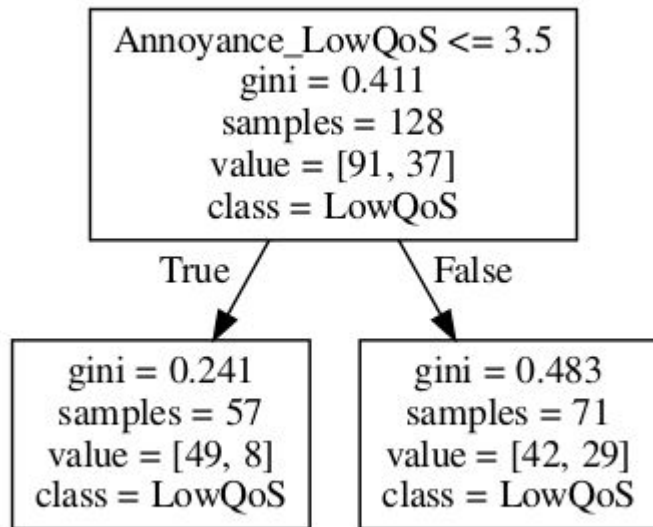


Balanced

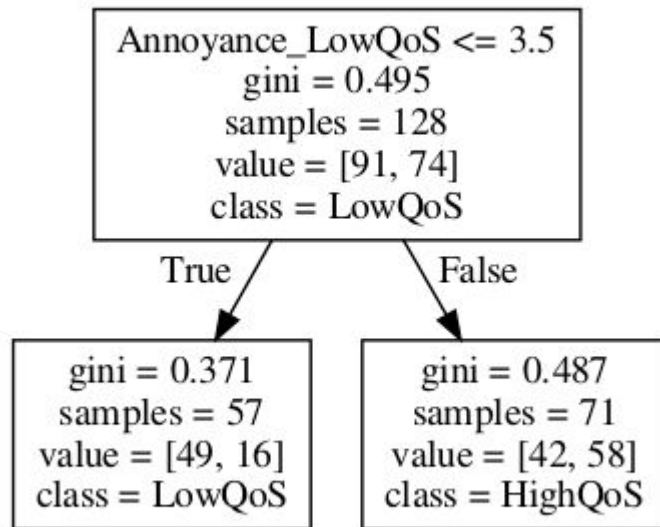


Tree Dif = 2

Unbalanced(No contribution)

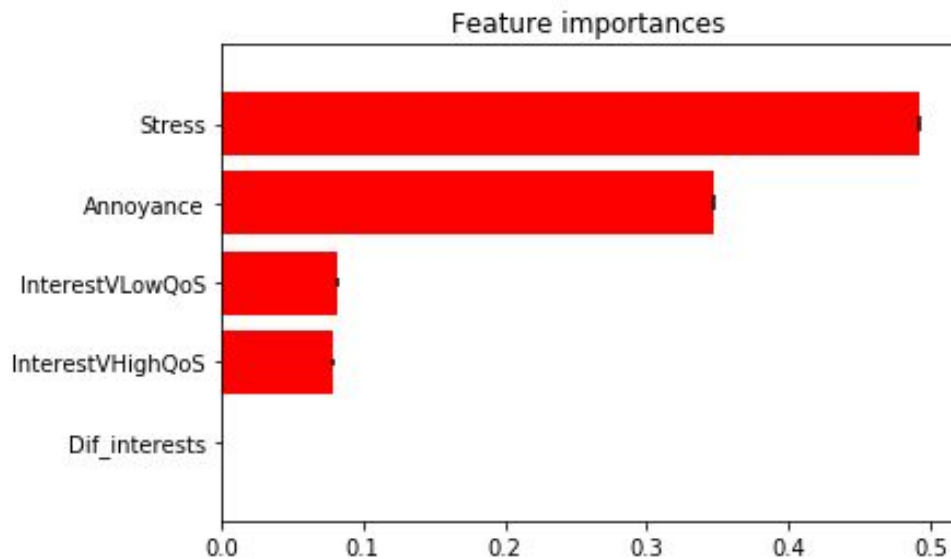


Balanced



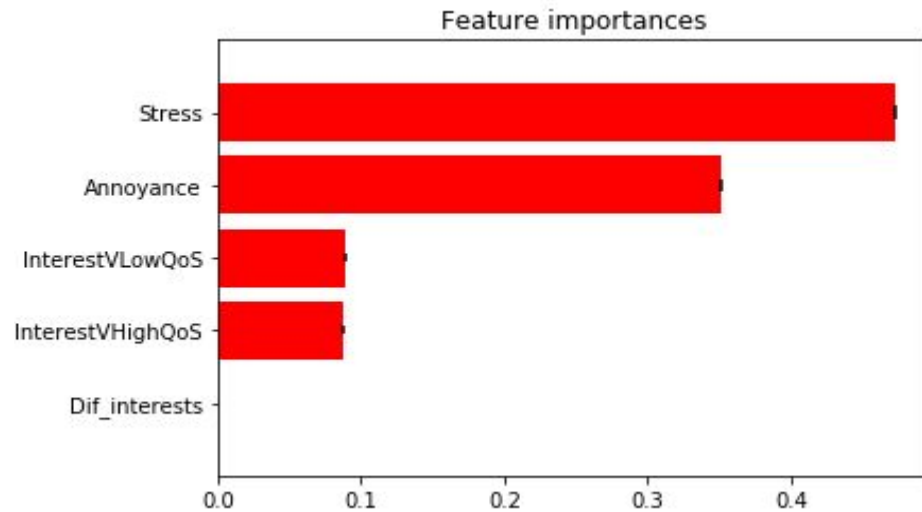
Random Forrest: Feature Importance (DiffInt = 0)

Accuracy: 83%



Extra Tree Dif = 1

Accuracy: 88%



Extra Tree Dif = 2

87% Accuracy

