UAM
Universidad Autónoma de Madrid

BACHELOR THESIS

Application of sentiment analysis on music recommendation

Student: Javier Wang Zhou
Advisor: Alejandro Bellogín Kouki

September 2023

Motivation

Universidad Autónoma de Madrid

Artists, Genres...

Emotions

Introduction: Recommender systems

Universidad Autónoma de Madrid

General

- Collaborative filtering
- Content-based
- Non-personalized

Context-aware

- Context (time, location...)
- Content-based

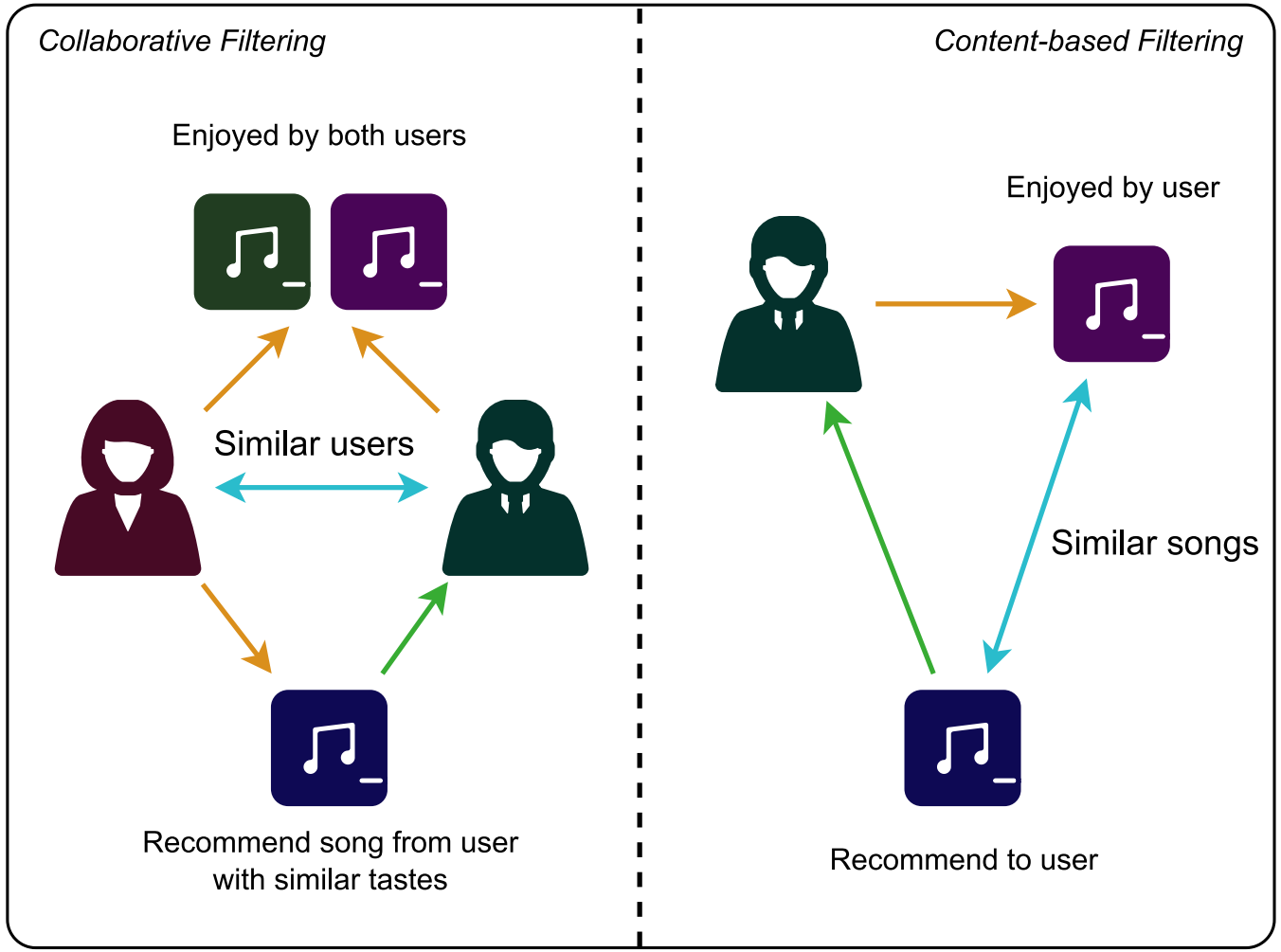
Sequential

- Temporal order of actions
- Next-item recommendation

Knowledge-based

- Knowledge graphs
- Semantic networks

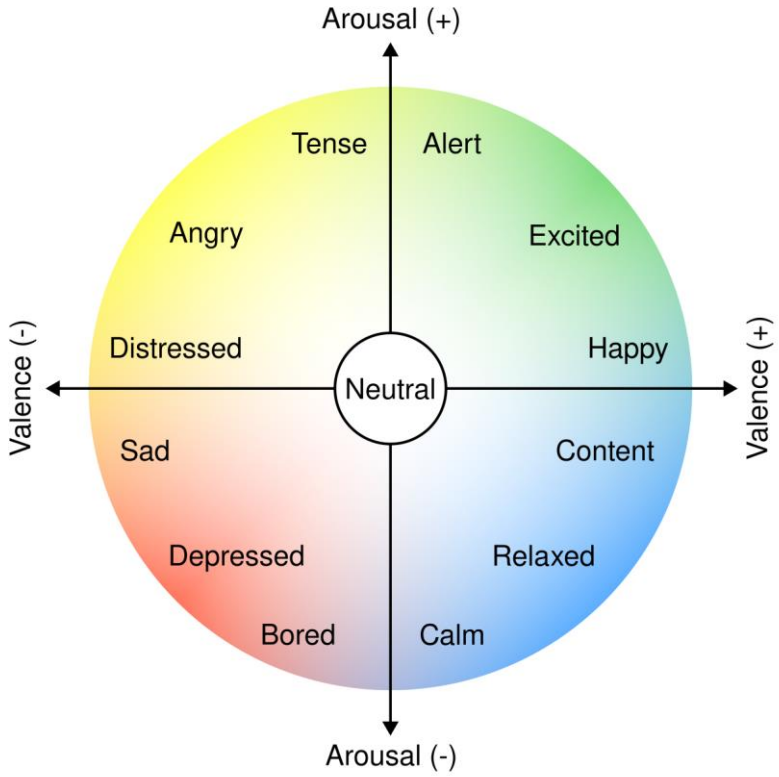
Introduction: Music recommendation

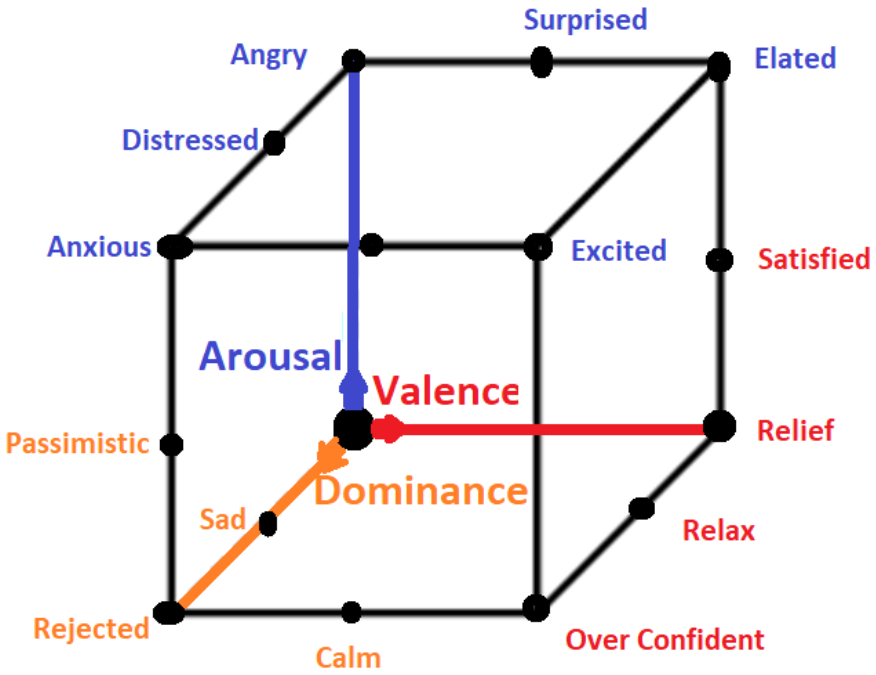


Introduction: Sentiment analysis

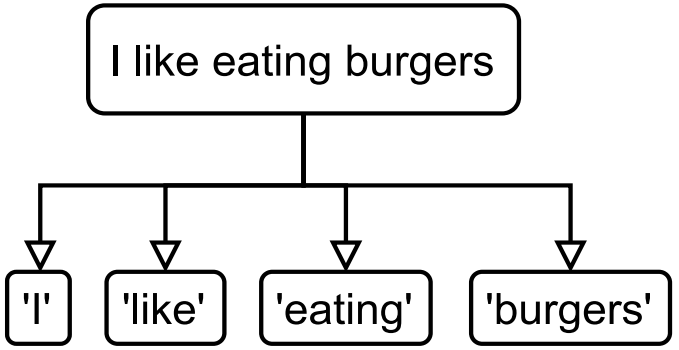
Universidad Autónoma de Madrid

- Models based on Valence / Arousal / Dominance.



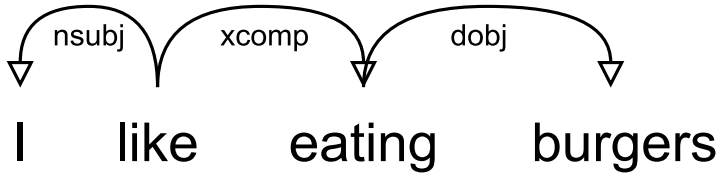


Introduction: Text processing

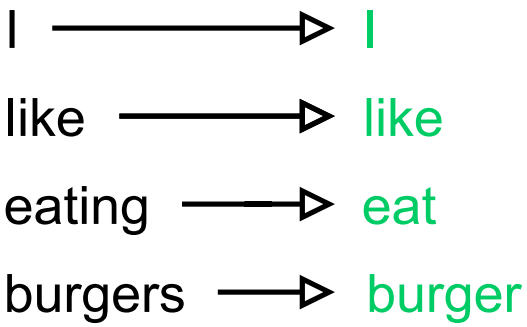


Tokenization

Part-of-speech Tagging

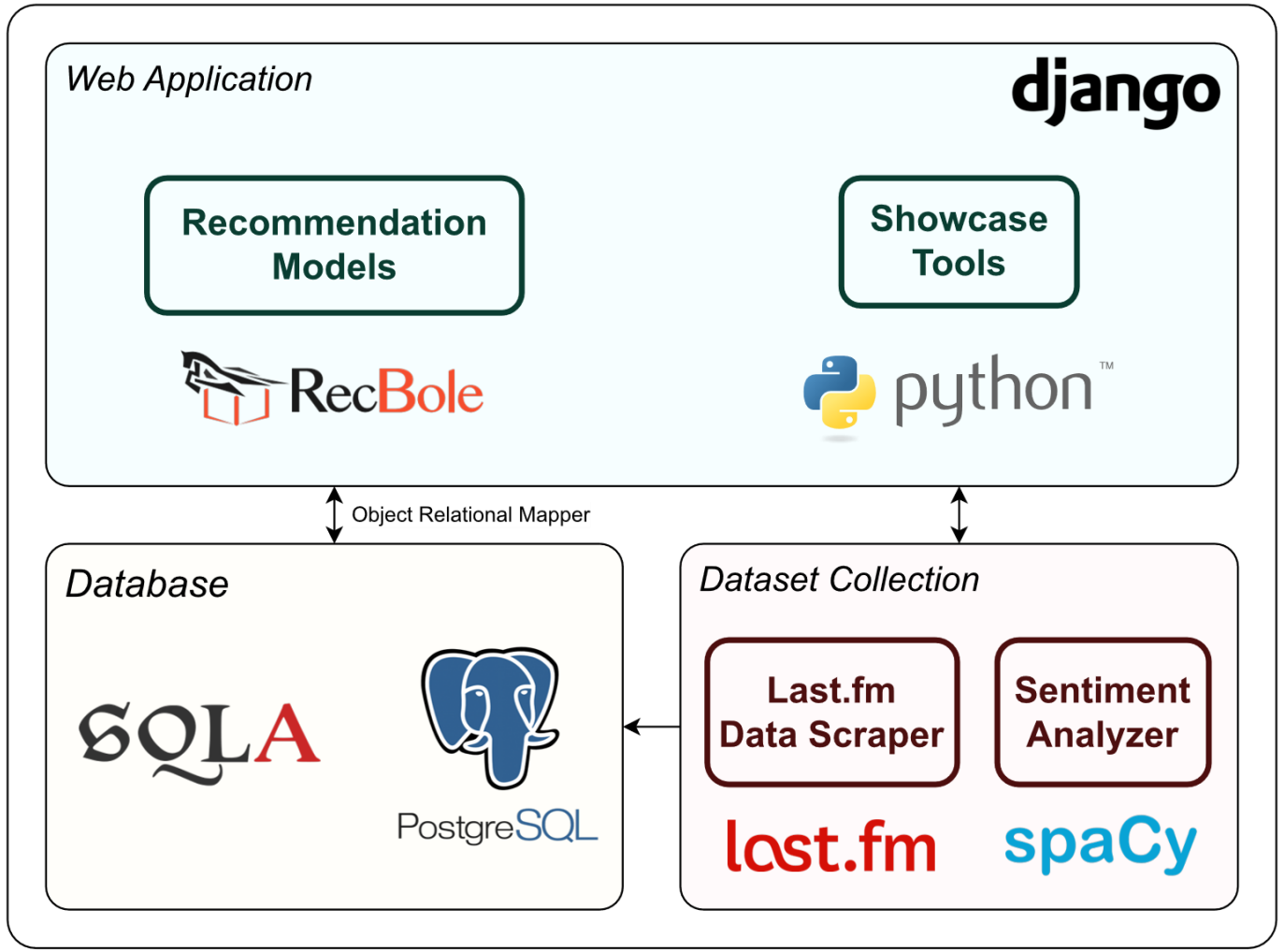


Dependency Parsing

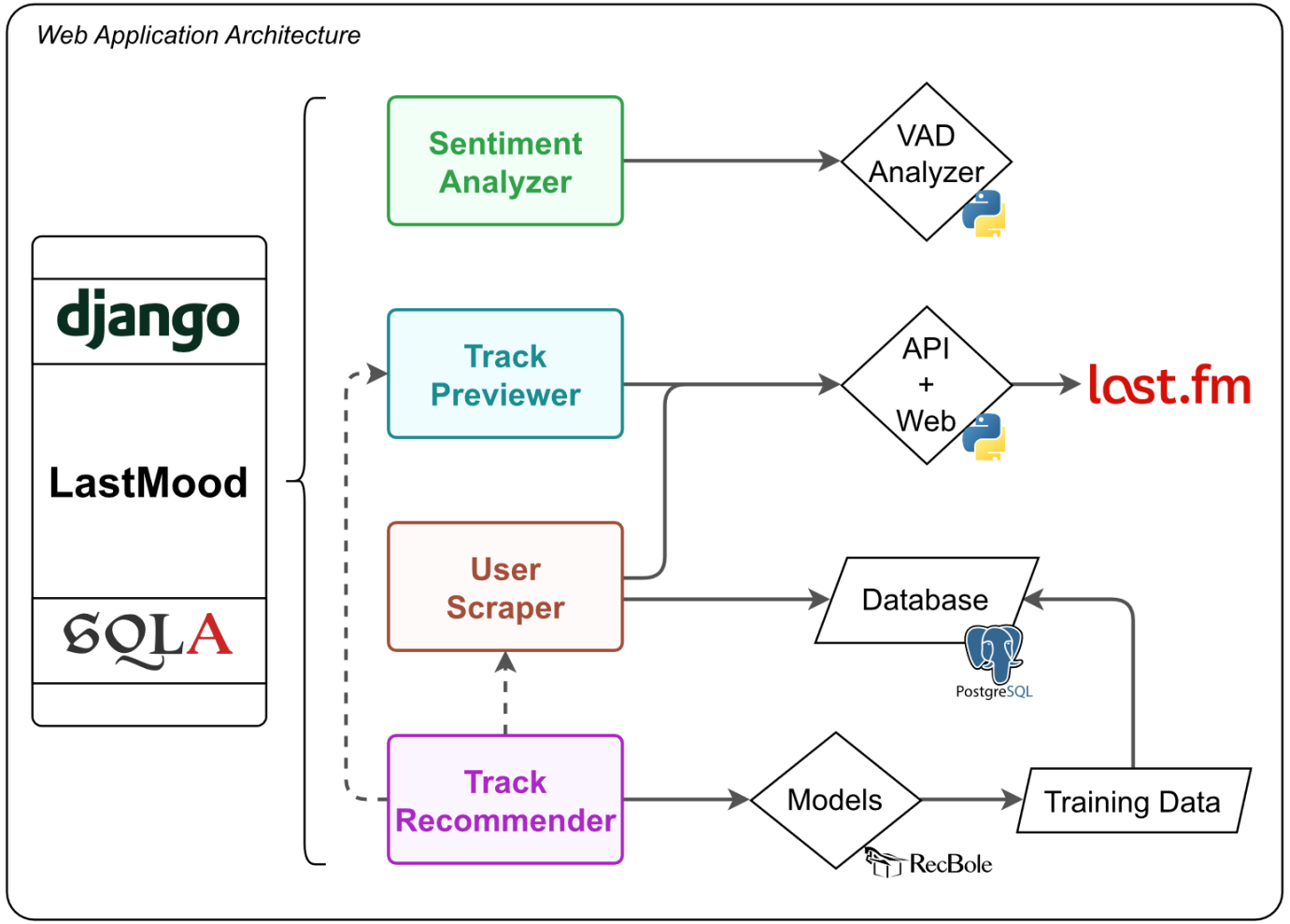


Lemmatization

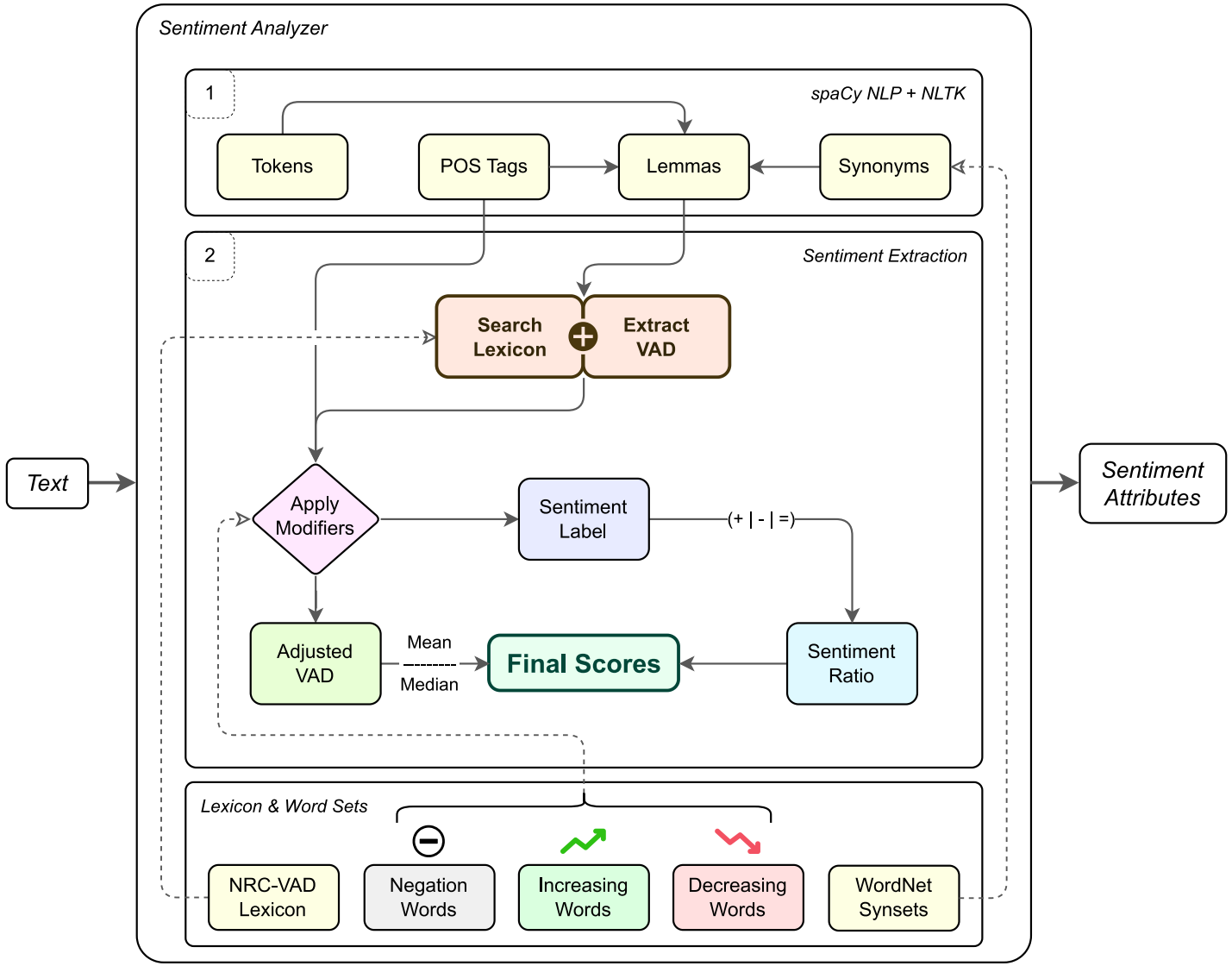
Project structure



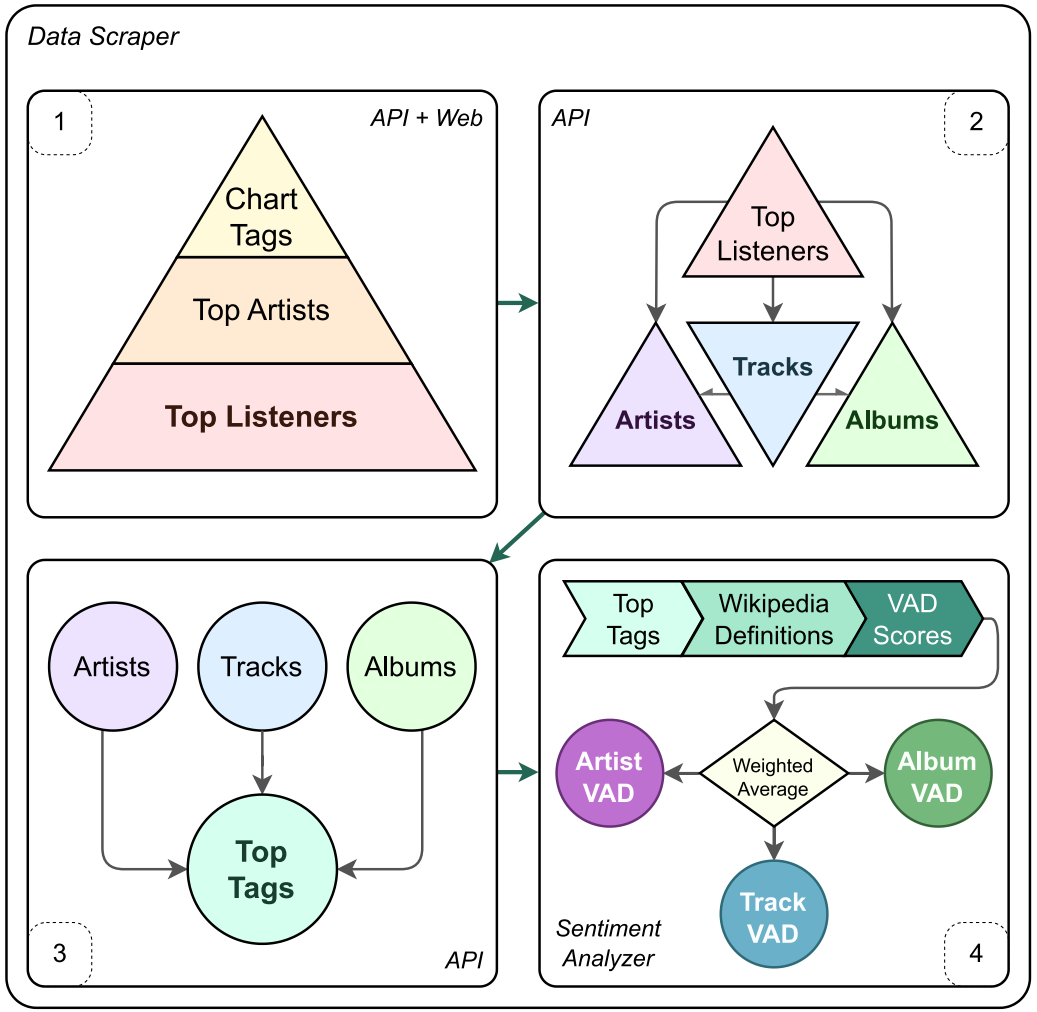
Web Application



Sentiment Analyzer



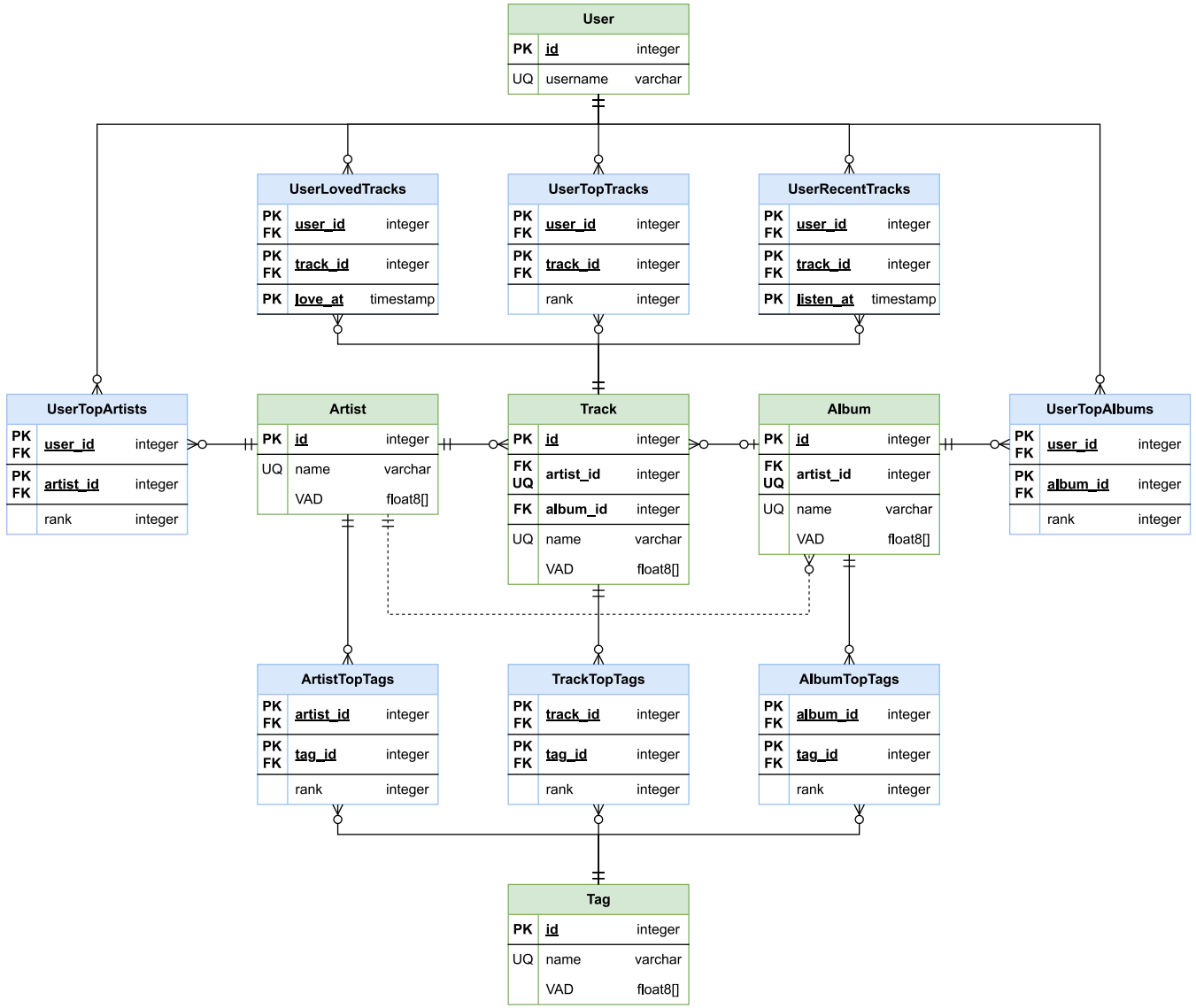
Data Scraper



1. Listeners
2. Items
3. Tags
4. Sentiment Values

- Sparsity?
- Objectivity?

Database



Universidad Autónoma de Madrid

Web Application: Sentiment Analyzer

Universidad Autónoma de Madrid

LastMood

Home
Track Previewer
Text VAD Analyzer
User Scraper
Recommendations
Login
Register

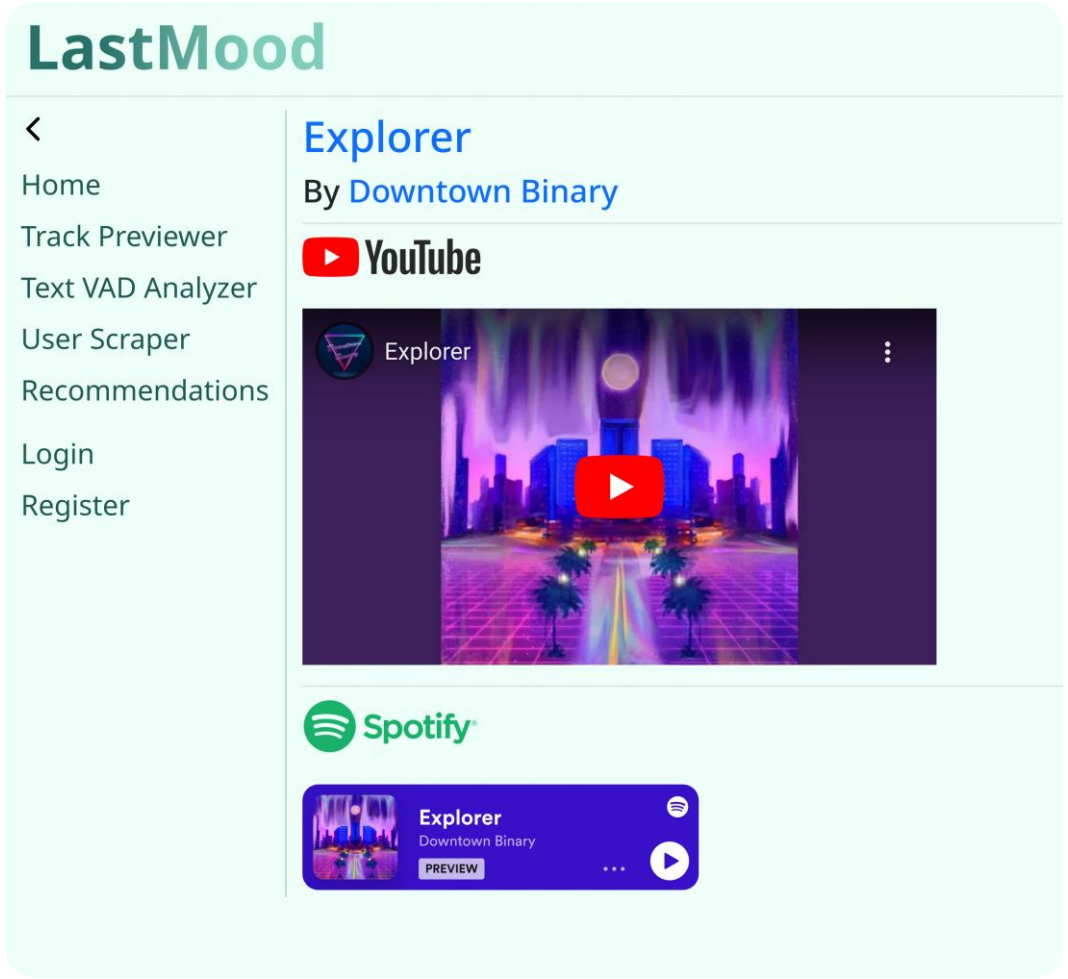
Text Sentiment Analysis Results

Sentence ID	Sentence	Valence	Sentiment Label
0	I really like eating burgers.	0.8908333333333333	positive
1	I don't like eating burgers.	0.2873333333333334	negative
2	I don't really hate burgers.	0.394625	negative

last.fm · Escuela Politécnica Superior · UAM Universidad Autónoma de Madrid

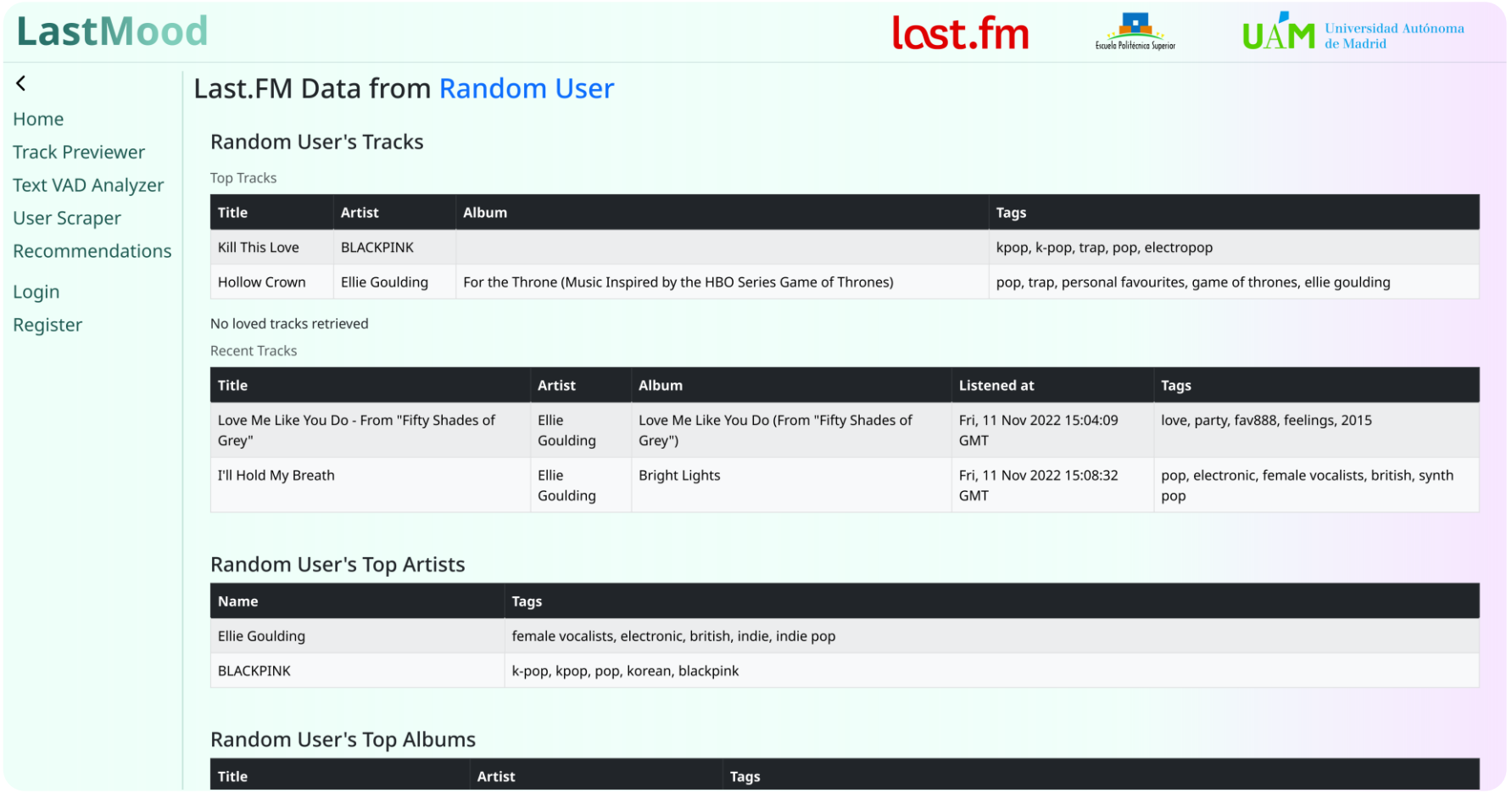
Sentiment Ratio	Arousal	Dominance	# Words Found	Found Words	All Words
1.0	0.4736666666666666	0.436	3 out of 3	['inc-like', 'inc-eat', 'inc-burger']	['like', 'eat', 'burger']
-1.0	0.5263333333333334	0.5640000000000001	3 out of 3	['neg-like', 'neg-eat', 'neg-burger']	['like', 'eat', 'burger']
0.0	0.685	0.41400000000000003	2 out of 2	['dec-hate', 'dec-burger']	['hate', 'burger']

Web Application: Track Previewer



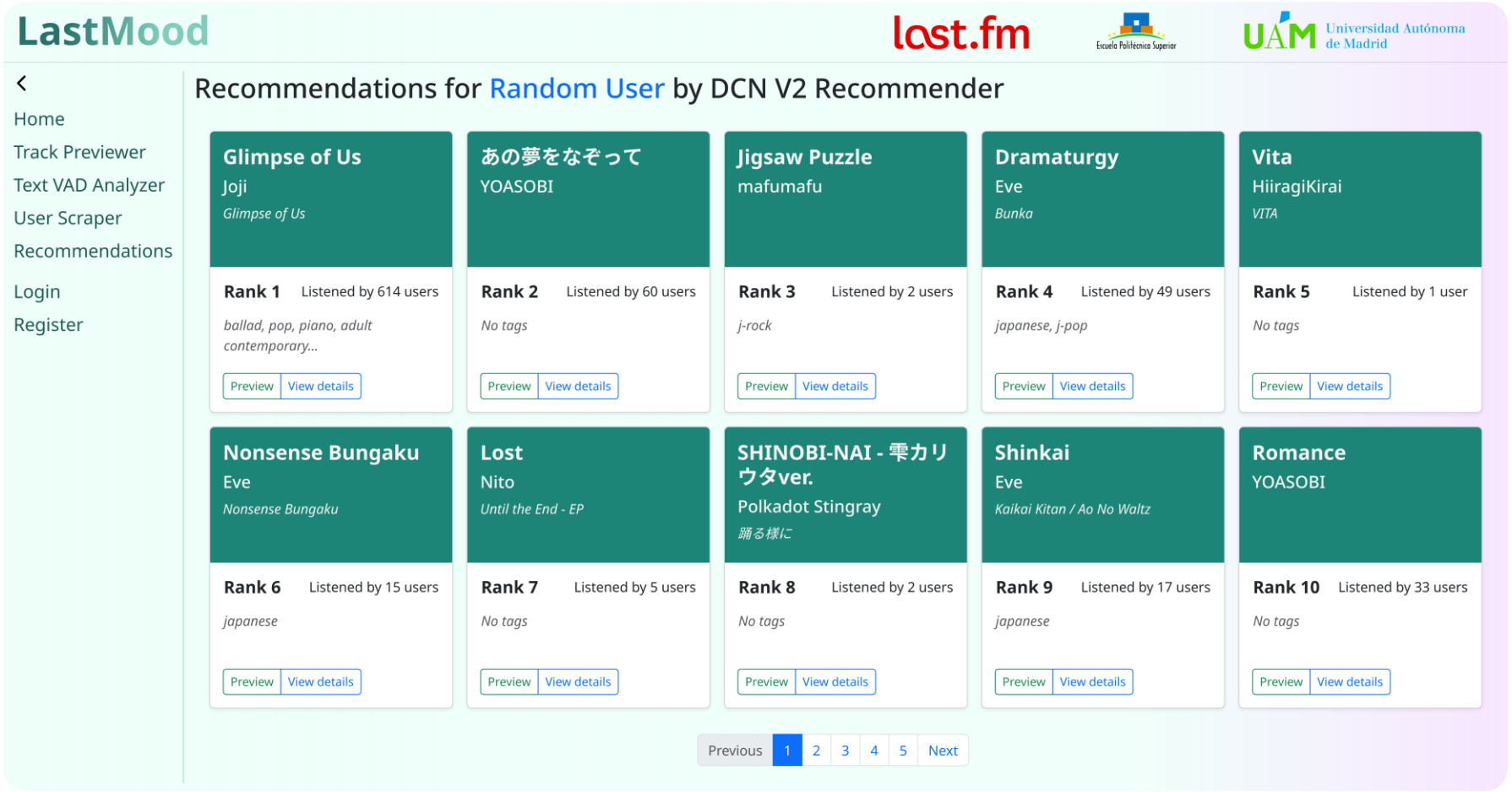
Universidad Autónoma de Madrid

Web Application: User Scraper

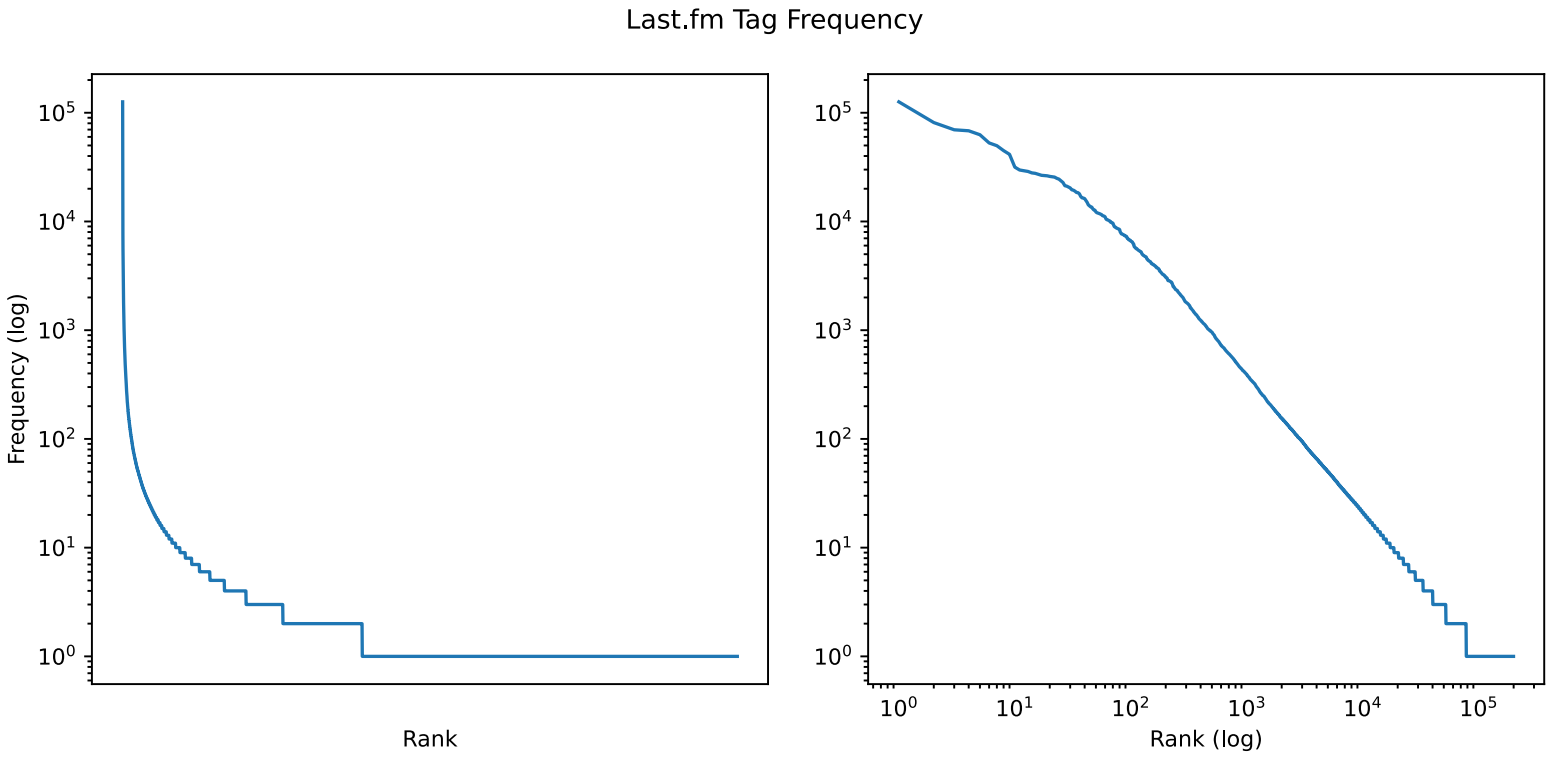


Web Application: Recommender

Universidad Autónoma de Madrid

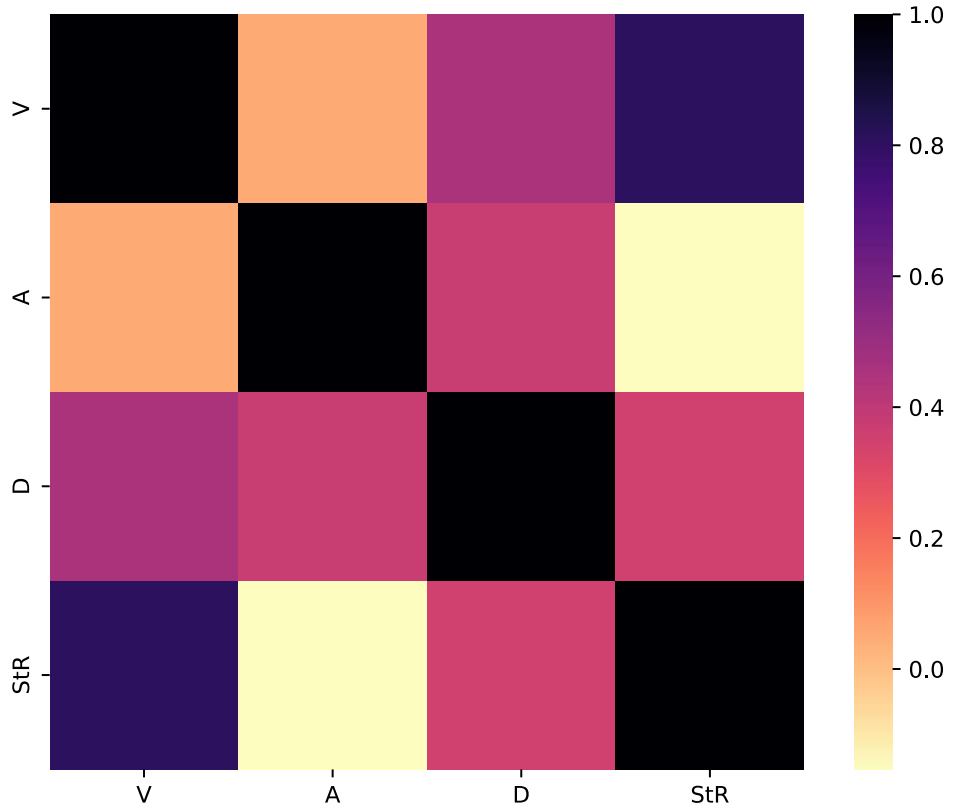


Data analysis: Tag frequency



- Zipf's Law
- Individual tag sparsity

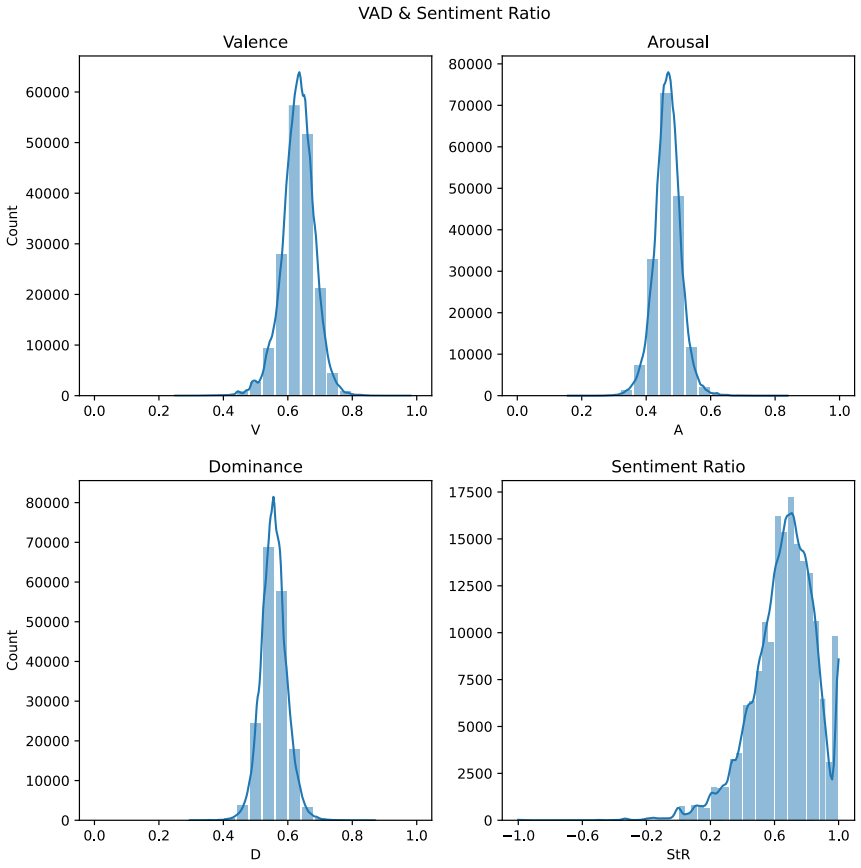
Data analysis: Sentiment correlations

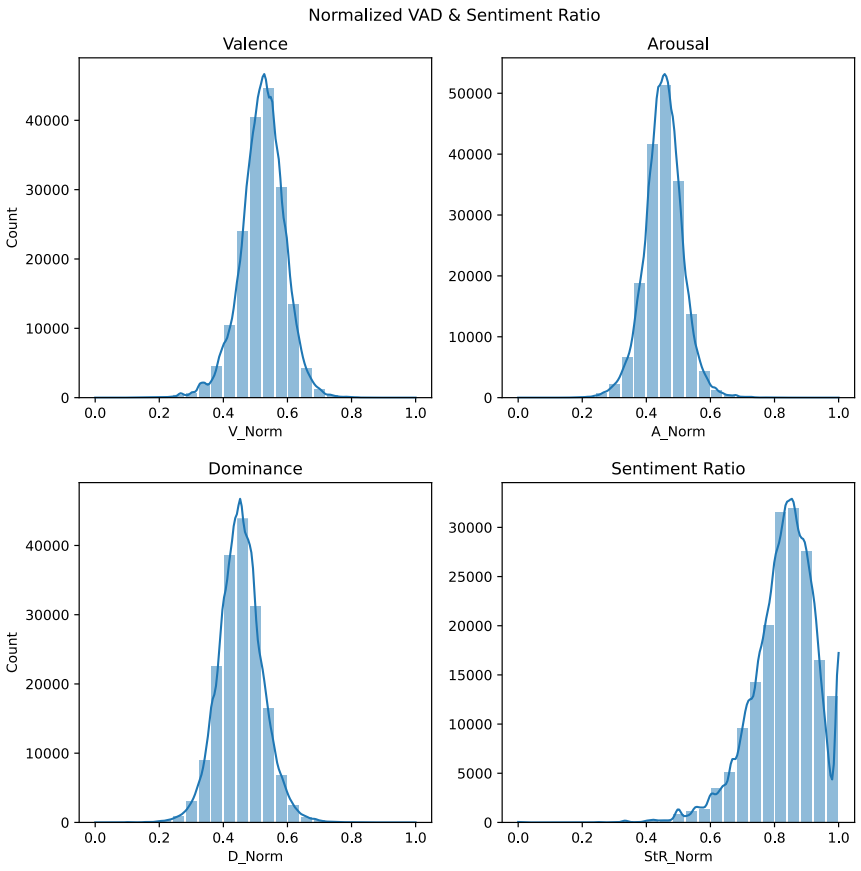


- Valence-Arousal
- St. Ratio-Arousal
- Valence-Dominance
- Arousal-Dominance
- St. Ratio-Dominance
- St. Ratio-Valence

V: Valence A: Arousal D: Dominance StR: Sentiment Ratio

Data analysis: Sentiment distributions

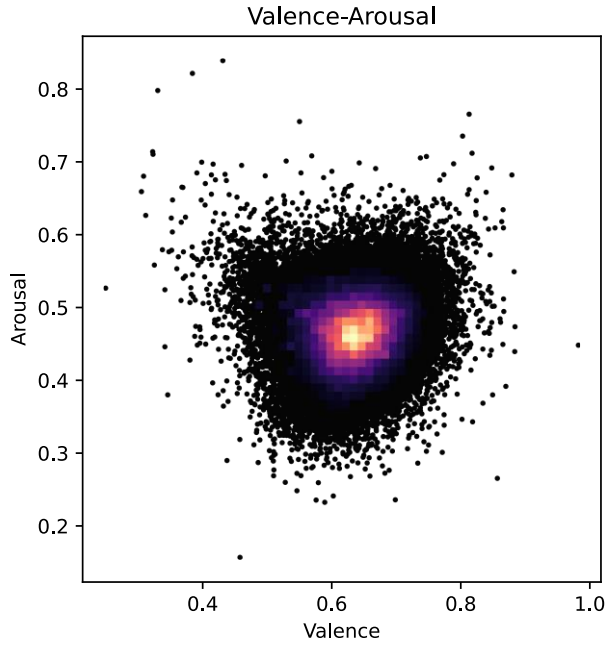


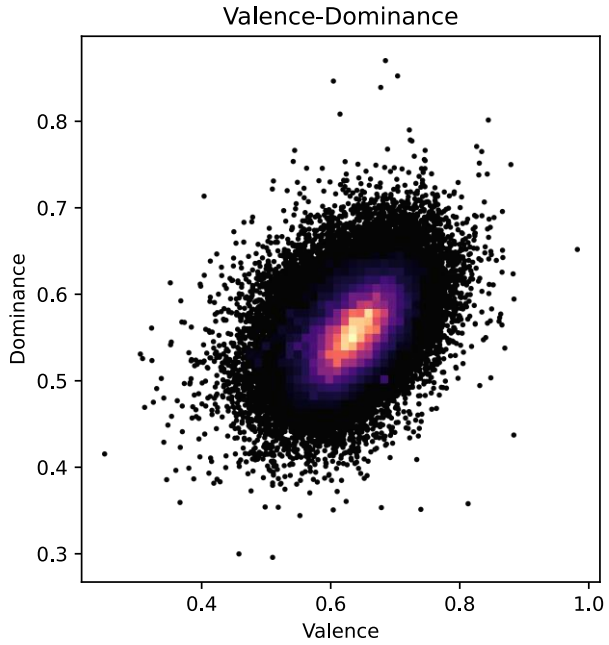


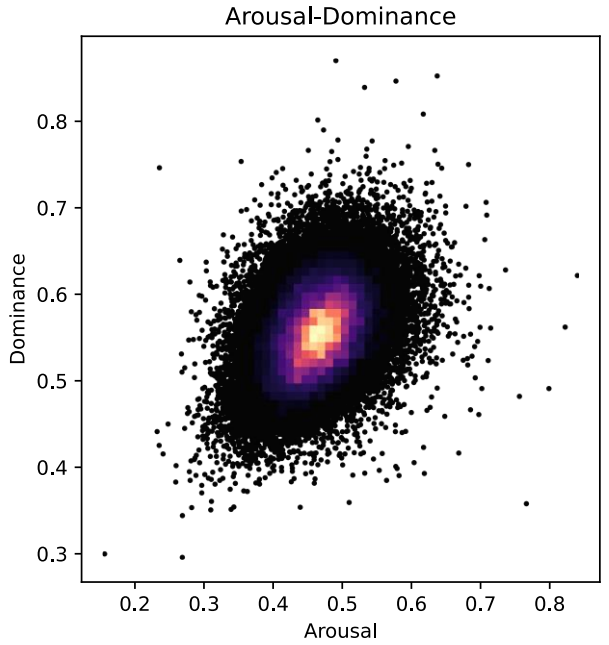
- Narrowly ranged values -> Normalization

Data analysis: Sentiment distributions

Universidad Autónoma de Madrid







V-shaped

Linear

Linear

Steeper

Testing: Environment

Universidad Autónoma de Madrid

Component	Specifications
CPU	AMD Ryzen 5 5600X
GPU	NVIDIA RTX 3060 Ti 8 GB VRAM
RAM	32 GB
O.S.	Windows 11
Python Environment	Python 3.9.16 (conda)

- Parallel work in separate environment

Testing: Data

- Preprocessing
- Feature selection
- Feature integration
- Testing subset

Testing: Recommendation libraries

Universidad Autónoma de Madrid

surprise

- 👍 Ease of use
- 👍 Dataset adaptability
- 👎 Limited models
- 👎 Documentation & customization

RecBole

- 👍 Many models
- 👍 Documentation & customization
- 👍 Optimal (GPU)
- 👍 Configuration with files
- 👎 Learning curve

Cornac

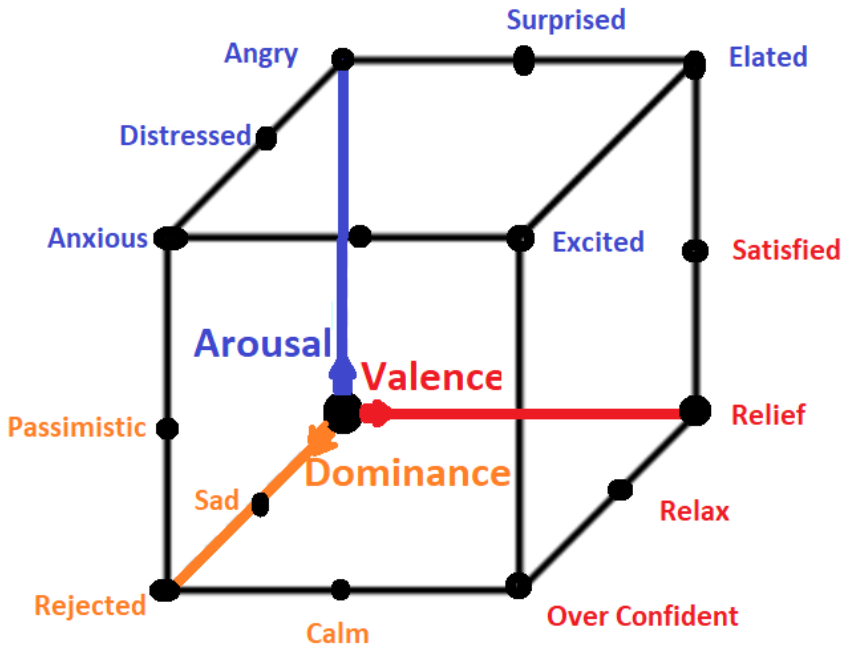
- 👍 Many models
- 👍 Documentation & customization
- 👎 Outdated

Testing: Models

Name	Type
Random	General
CosineSimilarity	General (Content-based)
Pop	General
ItemKNN	General
PNN	Context-aware
xDeepFM	Context-aware
DCN V2	Context-aware

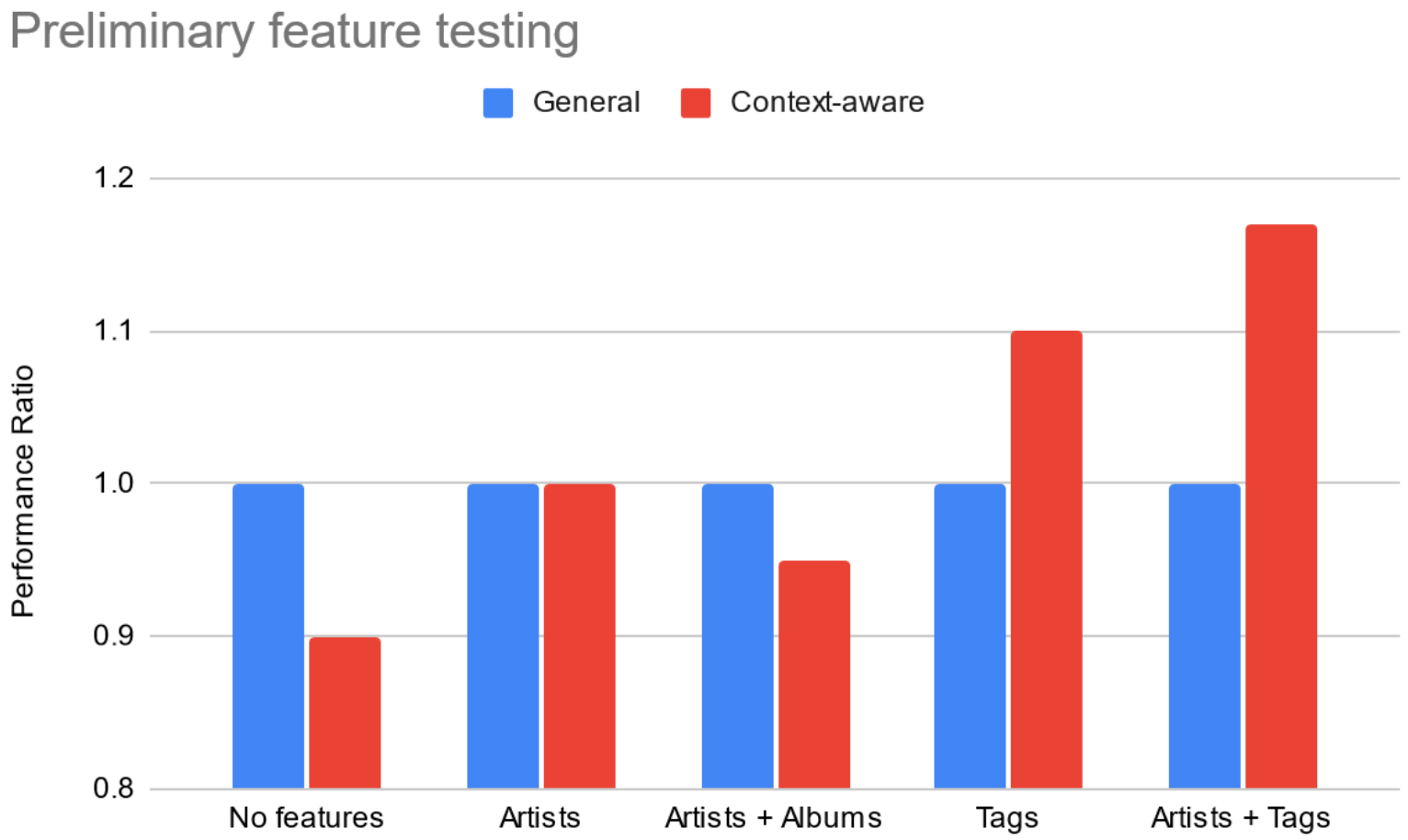
Results: Intuition

- Test with tags AND sentiment attributes



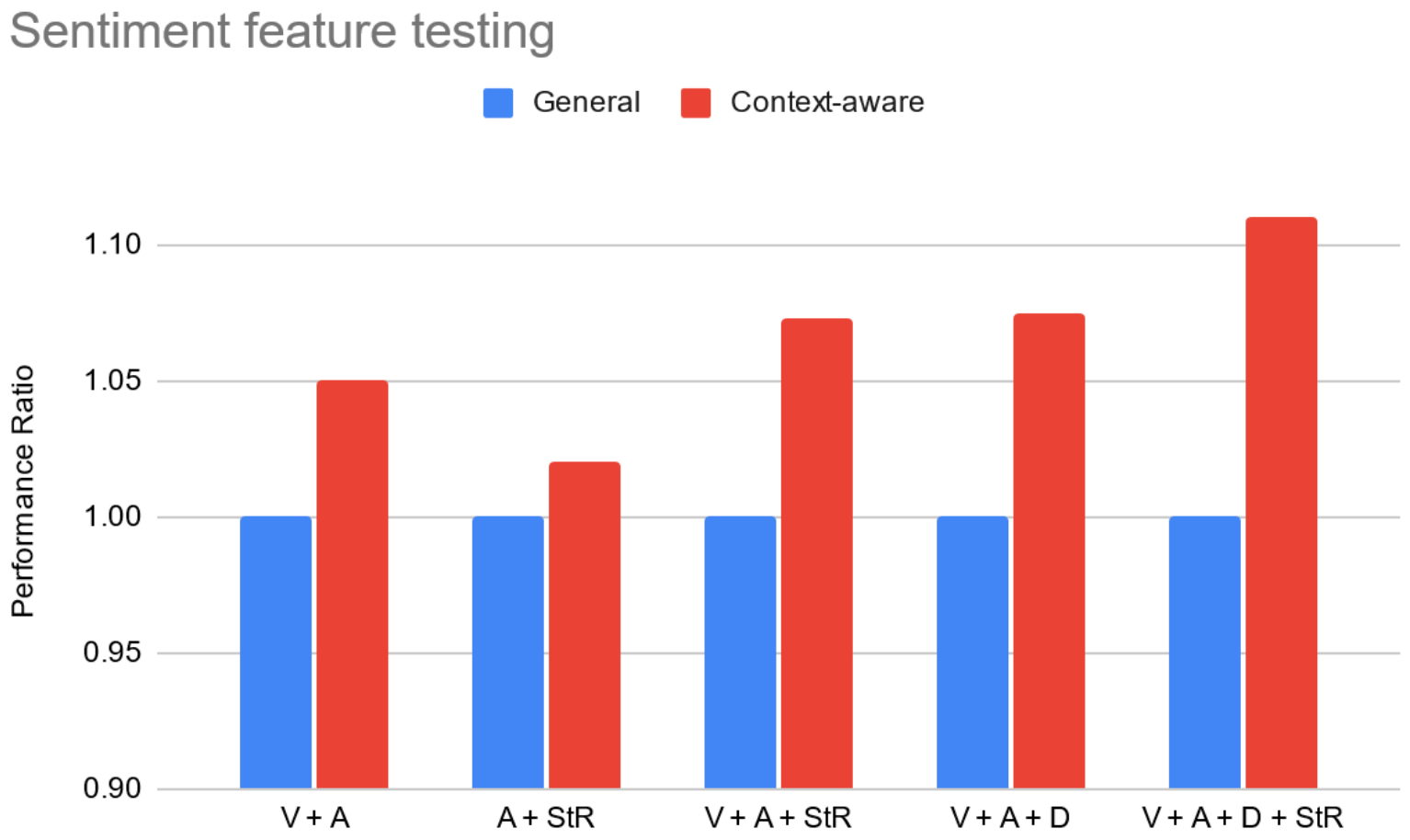
Results: Preliminary features

Universidad Autónoma de Madrid



Results: Sentiment features

Universidad Autónoma de Madrid



V: Valence A: Arousal D: Dominance StR: Sentiment Ratio

Results: Final features

Universidad Autónoma de Madrid

	Model	Testing Results @ 20									
		Preliminary (Artists + Tags)					Final (Artists + Tags + VAD + St.Ratio)				
		NDCG	Recall	Precision	mAP	MRR	NDCG	Recall	Precision	mAP	MRR
General	Random	0.03	0.06	0.01	0.01	0.04	=	=	=	=	=
General	CosineSimilarity	0.14	0.28	0.05	0.06	0.12	=	=	=	=	=
General	Pop	0.31	0.42	0.07	0.20	0.39	=	=	=	=	=
General	ItemKNN [35]	0.44	0.46	0.08	0.33	0.64	=	=	=	=	=
Context	PNN [36]	0.56	0.68	0.12	0.42	0.67	0.58	0.73	**0.13**	**0.44**	**0.67**
Context	xDeepFM [37]	0.57	0.67	0.12	**0.43**	**0.68**	**0.60**	**0.76**	**0.13**	**0.44**	**0.67**
Context	DCN V2 [38]	**0.58**	**0.71**	**0.13**	**0.43**	0.66	0.56	0.69	0.12	0.42	0.65

- Artists + Tags + VAD + Sentiment Ratio

Conclusions

Universidad Autónoma de Madrid

Artists, Genres...

Emotions

Future Work

Universidad Autónoma de Madrid

- Web application UX / performance testing
- Subjective texts for analysis (lyrics)
- Further research on VAD and other sentiment models

Thank you for your attention

Universidad Autónoma
de Madrid