

Jev and System One Models

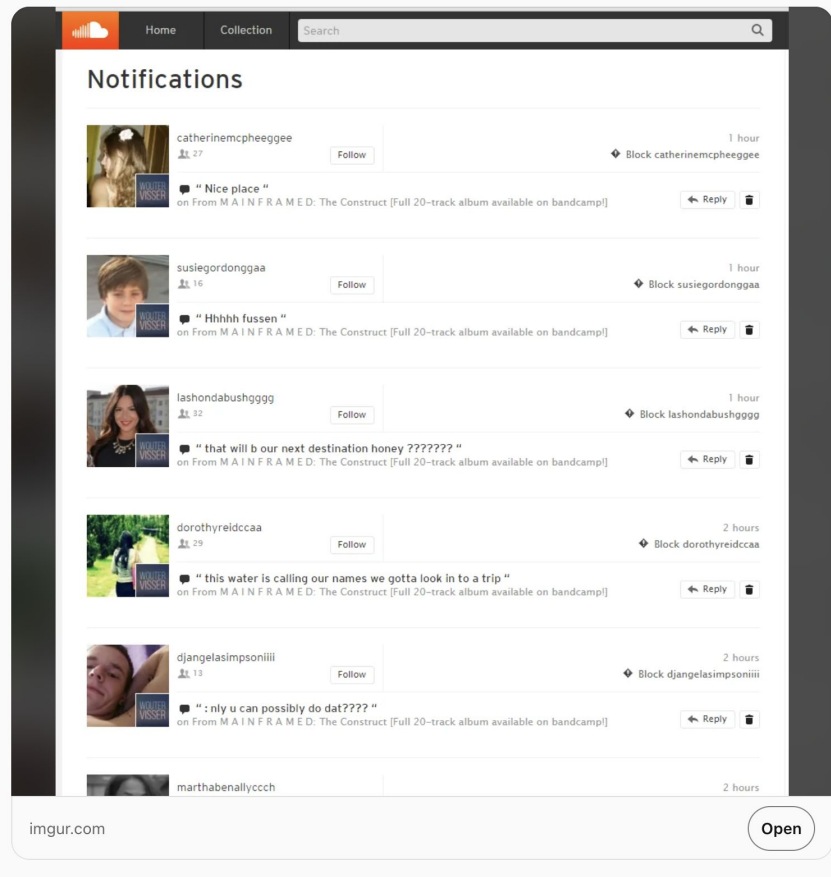
an engineer's perspective



r/WeAreTheMusicMakers · 11y ago
u/TheAmazingWJV



The spam on soundcloud is getting ridiculous. I feel like they're not even trying to stop it.



we were very much
trying...

... just failing
miserably



did what the smartest
guys in the room were
doing

Enron Email Dataset

This dataset was collected and prepared by the [CALO Project](#) (A Cognitive Assistant that Learns and Organizes). It contains data from about 150 users, mostly senior management of Enron, organized into folders. The corpus contains a total of about 0.5M messages. This data was originally [made public, and posted to the web](#), by the [Federal Energy Regulatory Commission](#) during its investigation.

*this is what passed as
big data circa 2011*

 **Out Tha Door 2014-04-26**

59 Comments In playlists In groups Recommended

< Back

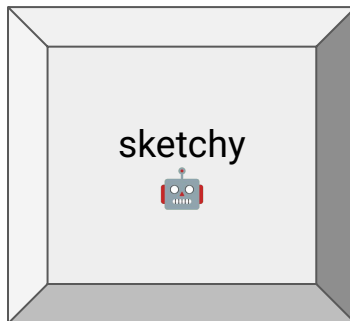
 You [Jeff3] took out Alicia's highest 'walk' here. thought it was too epic too early. more o this way at 0.56

 You [Jeff] backing 'I'm gonna walk' from Alicia is too thin at 4.00

 You @assemblethenoise: [Jeff3] jatworlded her some more (added takes) and adde as well. at 4.00

 You [JEFF2] I'd layer Alicia's ooh ahh in here but under Scott. It's too naked as is at 2.35

 You @assemblethenoise: [Jeff3] did it. at 2.35



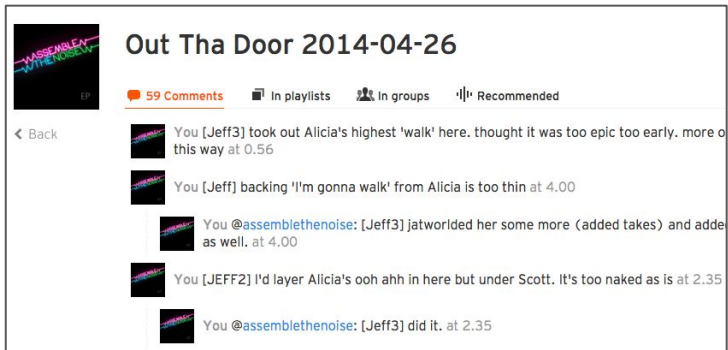
```
{  
  is_spam: 0.78,  
  is_not_spam: 0.1  
}
```

**promising results after
~1.5 year and 1 PhD
and 2 engineers**

— how do we
improve further?

— need another ~year,
and a lot more
relevant, manually
labeled, data

omg, general purpose classifiers?!?!



Out Tha Door 2014-04-26

59 Comments In playlists In groups Recommended

< Back

- You [Jeff3] took out Alicia's highest 'walk' here, thought it was too epic too early, more o this way at 0:56
- You [Jeff] backing 'I'm gonna walk' from Alicia is too thin at 4:00
- You @assemblethenoise: [Jeff3] jatworlded her some more (added takes) and adde as well, at 4:00
- You [JEFF2] I'd layer Alicia's ooh ahh in here but under Scott. It's too naked as is at 2:35
- You @assemblethenoise: [Jeff3] did it, at 2:35



```
{  
  is_spam: 0.78,  
  is_not_spam: 0.1  
}
```

except

- too slow
- too expensive
- too much hallucination

it can be done, but needs a lot of tricks

Phil Calçado Contact About Publications & Talks

Building AI Products—Part II: Task-Oriented vs. Component-Oriented Pipelines

Mar 14, 2025

• AI • Software Engineering • Agents •

Task-Oriented vs. Component-Oriented Pipelines

A while back, I kicked off a three-part series on how we built an AI-powered Chief of Staff for

<https://philcalcado.com/2025/03/14/building-ai-products-part-ii.html>

omg, general purpose classifiers?!?!



Out Tha Door 2014-04-26

59 Comments In playlists In groups Recommended

< Back

- You [Jeff3] took out Alicia's highest 'walk' here, thought it was too epic too early, more o this way at 0.56
- You [Jeff] backing 'I'm gonna walk' from Alicia is too thin at 4.00
- You @assemblethenoise: [Jeff3] jatworlded her some more (added takes) and added as well, at 4.00
- You [JEFF2] I'd layer Alicia's ooh ahh in here but under Scott. It's too naked as is at 2.35
- You @assemblethenoise: [Jeff3] did it, at 2.35



```
{  
  is_spam: 0.78,  
  is_not_spam: 0.1  
}
```

let's try one-shotting this

SPAM CLASSIFICATION · 100 LABELED MESSAGES

25 messages / request

Results

ngl this ain't great

	Jev	GPT-4.1 mini
Correct classifications	67 / 100	89 / 100
Spam caught	22 / 50	40 / 50
False alarms	5	1
Spam detection balance (F1)	57.1%	87.9%
Run time	1.79 seconds	12.97 seconds
Estimated API cost	\$0.0016	\$0.0082

Hugging Face

[Models](#) [Datasets](#) [Spaces](#)

mshenoda / **roberta-spam**

Like 33

Text Classification

Transformers

PyTorch

Safetensors

mshenoda/spam-messages

English

roberta

Model card

Files and versions

xet

Community 3

RoBERTa based Spam Message Detection

Spam messages frequently carry malicious links or phishing attempts posing significant threats to both organizations and their users. By choosing our RoBERTa-based spam message detection system, organizations can greatly enhance their security infrastructure. Our system effectively detects and filters out spam messages, adding an extra layer of security that safeguards organizations against potential financial losses, legal consequences, and reputational harm.

Found this model useful:

Your feedback is important and would help keep this relevant.

Metrics

Loss Train / Validation	Accuracy(99.06%) Validation	Precision(99.71%) / Recall(99.34%) Validation	Confusion Matrix Testing Set
---------------------------------------	---	---	--

would you look at that

mshenoda

[Docs](#) [API Console](#) [Menu](#)

Sep 11, 2026

Lies, Damned Lies, And Benchmarks

everybody is cheating
everything is in-distribution

real-world AI requires data engineering

1. classified each liked track by similar artist/genre/vibe
2. clustered and consolidated
3. identified patterns within each group
4. extracted propositions for each cluster
5. merged overlapping propositions across all clusters
6. user the normalized propositions list to create criteria questions for jev

```

{
  "0:indie_rock_and_folk": {
    "type": "noul",
    "instructions": "Does `candidates[0]` fit this listener affinity: Listener has a strong preference for indie rock and indie folk music characterized by emotionally expressive songwriting and guitar-based arrangements?",
    "criteria": {
      "true": {
        "what": "Artists such as The National, The Strokes, Angus & Julia Stone, Blind Pilot, Puzzle Muteson, and genres including indie rock, indie folk, baroque pop, and singer-songwriter.",
        "not_for": "Generic popularity, vague similarity to rock, or missing metadata about indie or folk style are insufficient."
      },
      "false": {
        "what": "No evidence for a liking of indie rock or indie folk styles."
      }
    }
  },
  "0:classical_and_opera": {
    "type": "noul",
    "instructions": "Does `candidates[0]` fit this listener affinity: Listener has an affinity for classical music including operatic and orchestral works, spanning traditional and contemporary classical styles?",
    "criteria": {
      "true": {
        "what": "Genres such as classical, opera, contemporary classical, with composers like Wagner, Mahler, Beethoven, and renowned orchestras and conductors.",
        "not_for": "Casual or popular classical crossover without serious orchestral or operatic components is insufficient."
      },
      "false": {
        "what": "No evidence for classical or opera appreciation."
      }
    }
  }
}

```

↓ Jump to latest message · Ctrl+End

500 tracks · 250 liked / 250 controls
25 tracks / request

MUSIC PREFERENCE · IN-DISTRIBUTION VS. OUT-OF-DISTRIBUTION

Results

	Jev	GPT-4.1 mini
Correct classifications	91.4%	98.8%
Known likes recognized	82.8%	97.6%
Out-of-profile controls rejected	100.0%	100.0%
False recommendations	0	0
Missed liked tracks	43	6
Run time	8.66 seconds	50.54 seconds
Estimated API cost	\$0.0338	\$0.0333

wdyt?