
Grocery Apps, ChatGPT, Sentiment Analysis

Ethan Hong





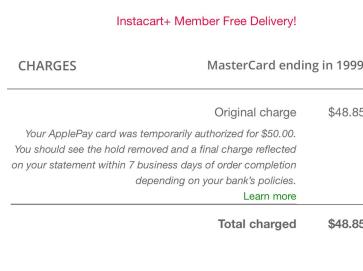
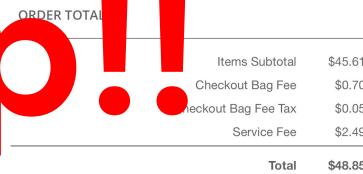
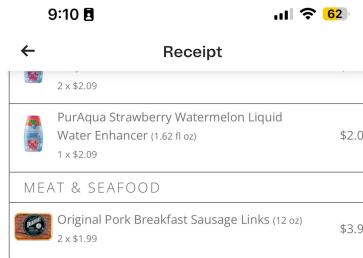
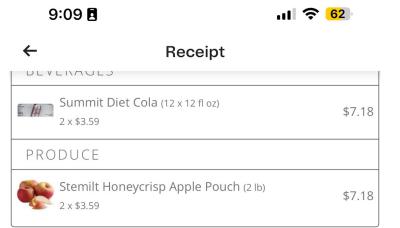
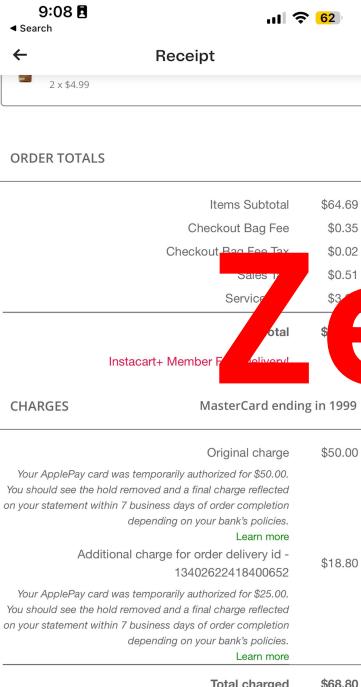
“To what extent can ChatGPT assist Grocery Delivery Apps through Sentiment Analysis in comparison to other Algorithms”

How I Arrived Here



Introduction

Zero Tip!!



Instacart shoppers say they face unforgiving metrics: 'It's a very easy job to lose'



Kaitlin Myers, an Instacart shopper, at a Whole

RKFORCE MANAGEMENT

As workers crave a regular paycheck, could the gig economy's days be numbered?

Paola Peralta April 19, 2023, 4:16 p.m. EDT 3 Min Read



Instacart's harsh ratings system hurts grocery delivery people like me

Our livelihoods depend on a ratings system that must change.

By Ehud Sopher Mar 19, 2021, 8:30am EDT

US news

This article is more than 2 years old

'Tired of being stepped on': Instacart workers urge customers to delete app

'Shoppers' across the US cite falling pay rates and unsafe working conditions ahead of company's rumored IPO



An Instacart worker delivers groceries in Tucson, Arizona, in 2020. Photograph: Cheney Orr/Reuters

Gloria Oladipo

Tue 5 Oct 2021 03.30 EDT



ork City's minimum wage law for delivery workers
l by the courts

New York State Court's Appellate Divi
d legislation that almost triples the mi
o-based delivery workers to \$19.96

Fantozzi | Dec 01, 2023



New York City delivery
drivers will be paid much

06-28-2023 | TECH

Doordash just announced its biggest app update ever. Here's what's new

Workers can now choose to work under a specific rate rather than getting paid different rates for each task.

Delivery Apps Often Fail to Deliver



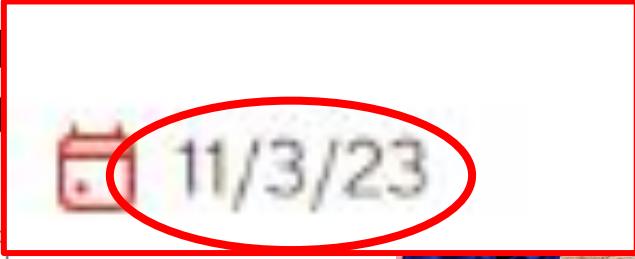
Instacart

Order Delivery
Groceries!
Food Delivery
Grocery Delivery
Pet Supplies
Services!

S · DEPARTMENT OF LABOR

January 9, 2024 at 8:06 AM EST

ing Updates to Our Ratings System



Nov 15, 2023

3 Possibilities

1. Done By Humans

2. Done By Another Algorithm

3. Done by ChatGPT

125,000x faster than the human neuron

And an overall accuracy of **92.1%** for counseling comments





—

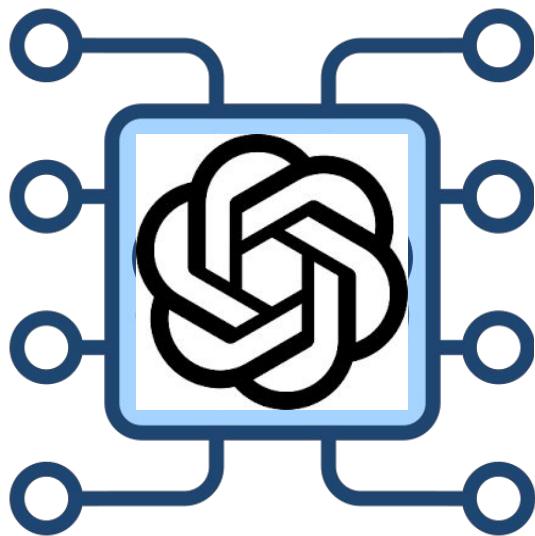
“To what extent can ChatGPT assist Grocery Delivery Apps through Sentiment Analysis in comparison to other Algorithms”

Method: Sentiment Analysis X Comparative Analysis

- 1. Create &
Conduct
Sentiment
Analysis**
- 2. Compare
Results**

My Goal

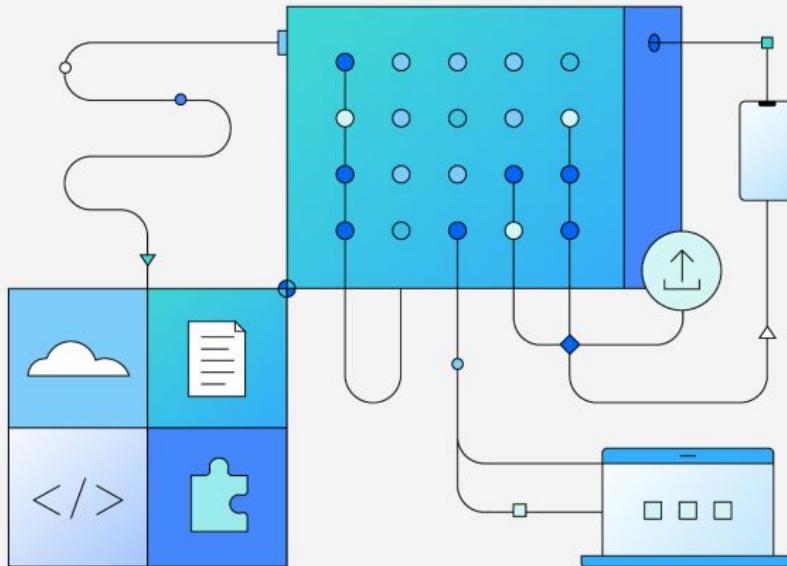
Overall Sentiment: [...]
Featured Issues: [...]



Hypothesis

“ChatGPT will outperform
the Top Sentiment Analysis
Algorithms in all metrics.”

Machine Learning



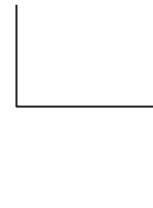
Lexicon/Dictionary Based

“Score closer to -1”



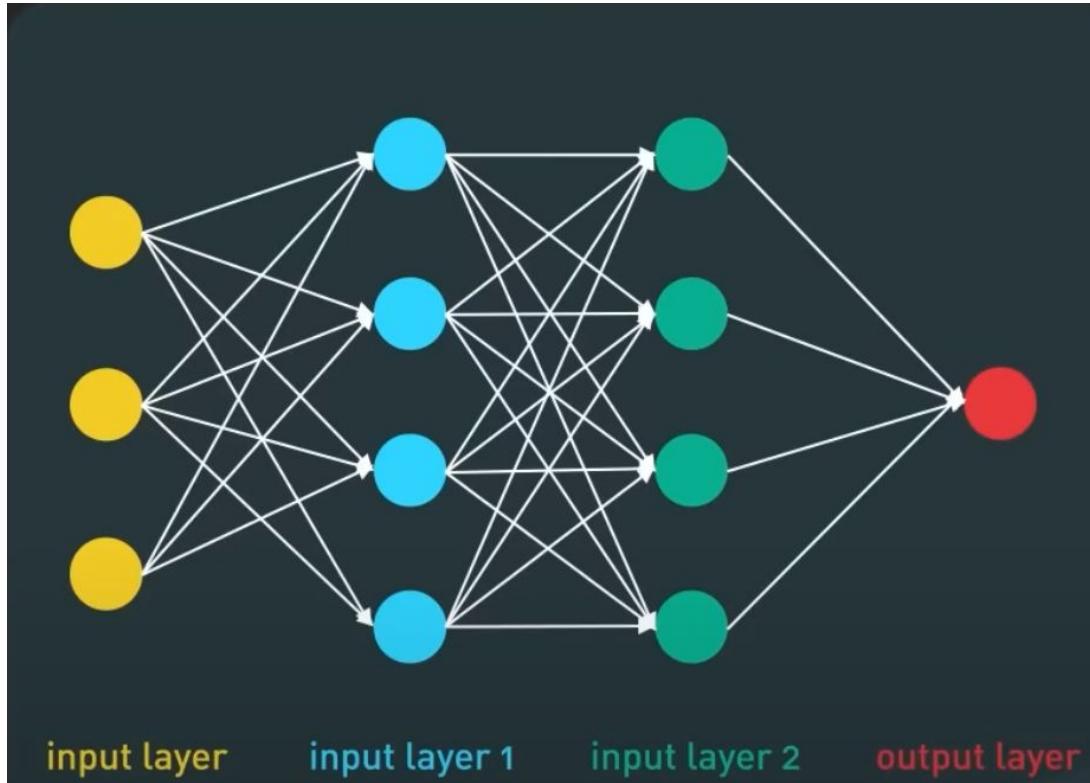
“disappointed”

“amazed”

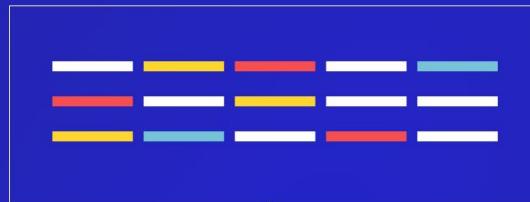


“Score closer to 1”

Deep Learning



Sentiment Analysis



SENTIMENT ANALYSIS



Literature Review!

Paper #1: A study on reviews of online grocery stores during COVID-19 pandemic using sentiment analysis

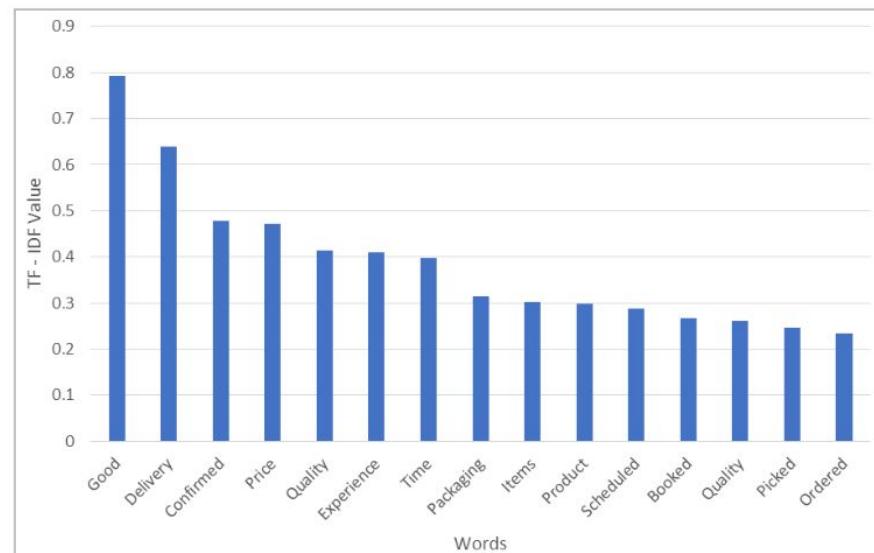
$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

$$F1 = \frac{2 \times precision \times recall}{precision + recall}$$

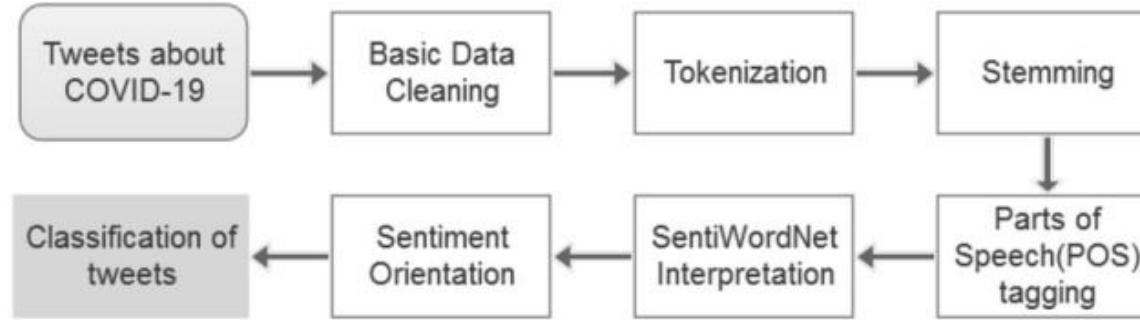
$$accuracy = \frac{TP + TN}{TP + FN + TN + FP}$$

Figure 1 Words with their TF-IDF values (see online version for colours)

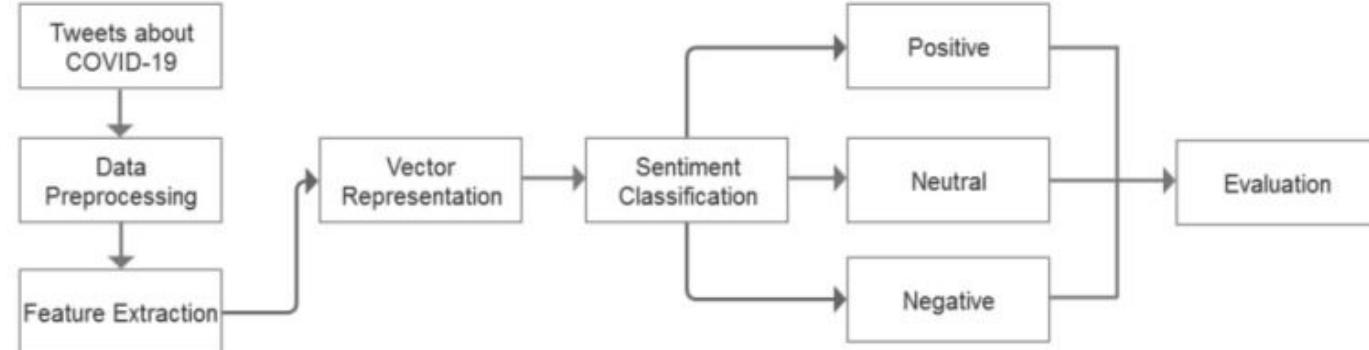


Paper #2: Sentiment analysis using Twitter data: a comparative application of lexicon- and machine-learning-based approach

Machine Learning Model:



Dictionary Based Model:



Paper #3: Customer reviews analytics on food delivery services in social media: a review

Table 3. Accuracy of AI algorithm used

Algorithm	Factors	Average Accuracy	Reference
Lexicon	Customer experiences, Quality control	87.33%	[46]-[48]
Natural language processing (NLP)	Quality control, Customer experiences	71.67%	[41], [50], [51]
Support vector machine (SVM)	Quality control	69.70%	[1], [7], [12], [39], [52], [53]
Text Mining	Quality control, Food quality, Service quality, Customer experience	67.94%	[12], [13], [29], [42]

Conclusion from Literature Review:

- There haven't been any studies done for ChatGPT from a Grocery Review Perspective
- ChatGPT's extraordinary Potential
- Understanding of a strategy towards this study





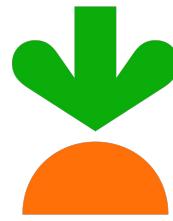
Method

Step #1: Gather Data/Reviews

MOST POPULAR



 DOORDASH



ConsumerAffairs



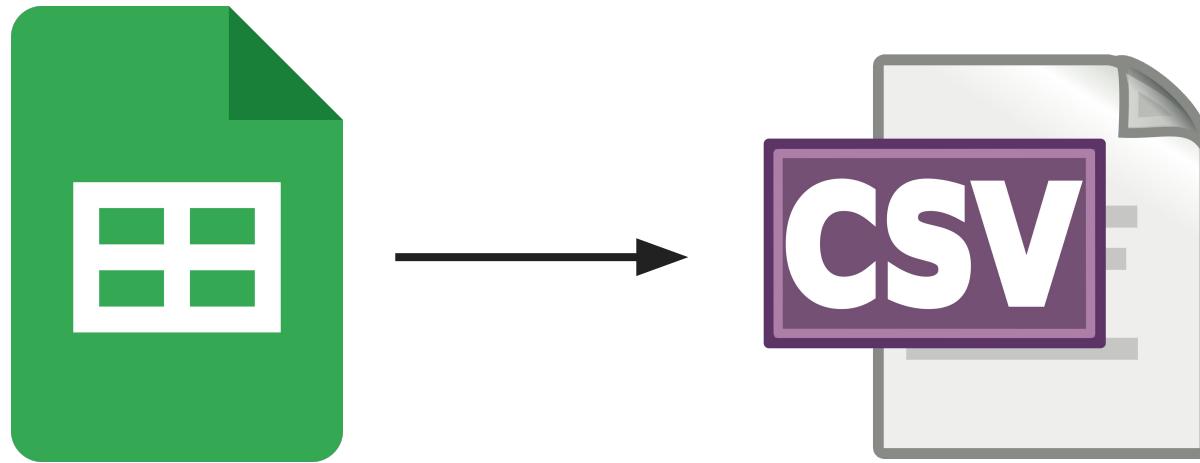
Least Popular



 Shipt

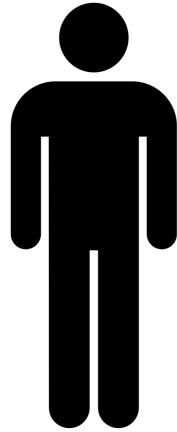


Step #2: Input into Google Sheets for CSV

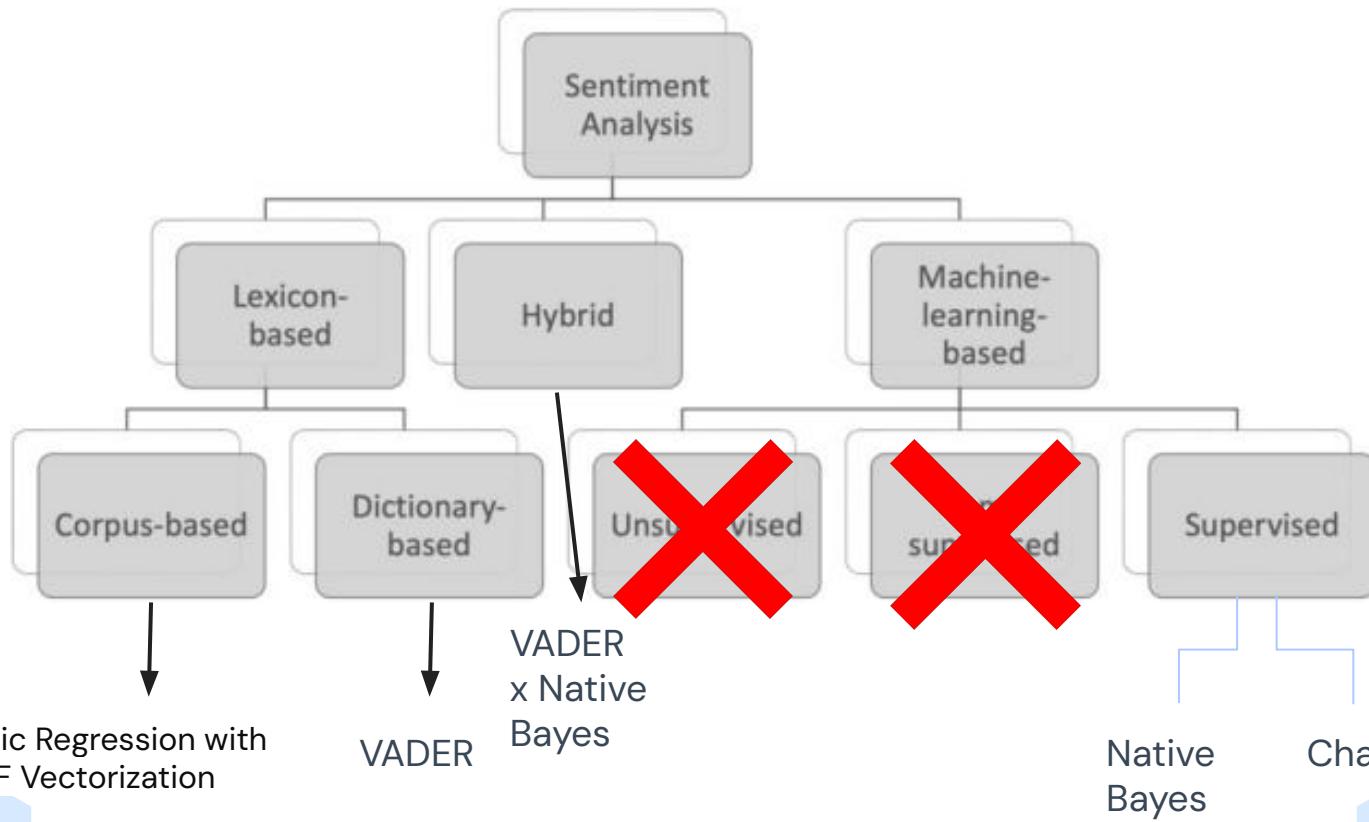


Step #3: Manual Sentiment Analysis

"Hmm..
Positive/Negative/Neutral"

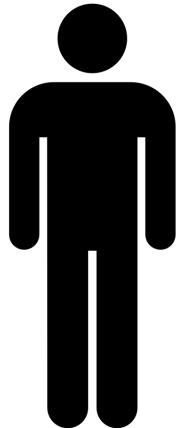


Step #4: Choose Algorithms



Step #5: Prompt ChatGPT

“Conduct a sentiment analysis of this subset of reviews and include your reasoning.”



“Ah yes. _____ Sentiment: ...”



Step #6: Preprocess & Clean Reviews

“.” HELLO → hello

Step #7: Create & Compare Sentiment Analysis Algorithms

Machine Learning (Native Bayes & ChatGPT)

1. Load Data
2. Imports
3. Preprocess Data
4. Vectorize Test Data
5. Train Model
6. Evaluate Model

Hybrid



Lexicon Based (VADER)

1. Load Data
2. Imports
3. Preprocess Data
4. Define the Dictionary
5. Determine Sentiments

Corpus- Based

1. Corpus Collection
2. Labeling:
3. Model Training
4. Model Evaluation
5. Model Deployment

Step#8: Conduct & Compare TF-IDF

TF-IDF Values of Reviews

TF-IDF Values of ChatGPT

Step #9: Error Analysis



Step #10: Overall Overview

1. Interpretability 

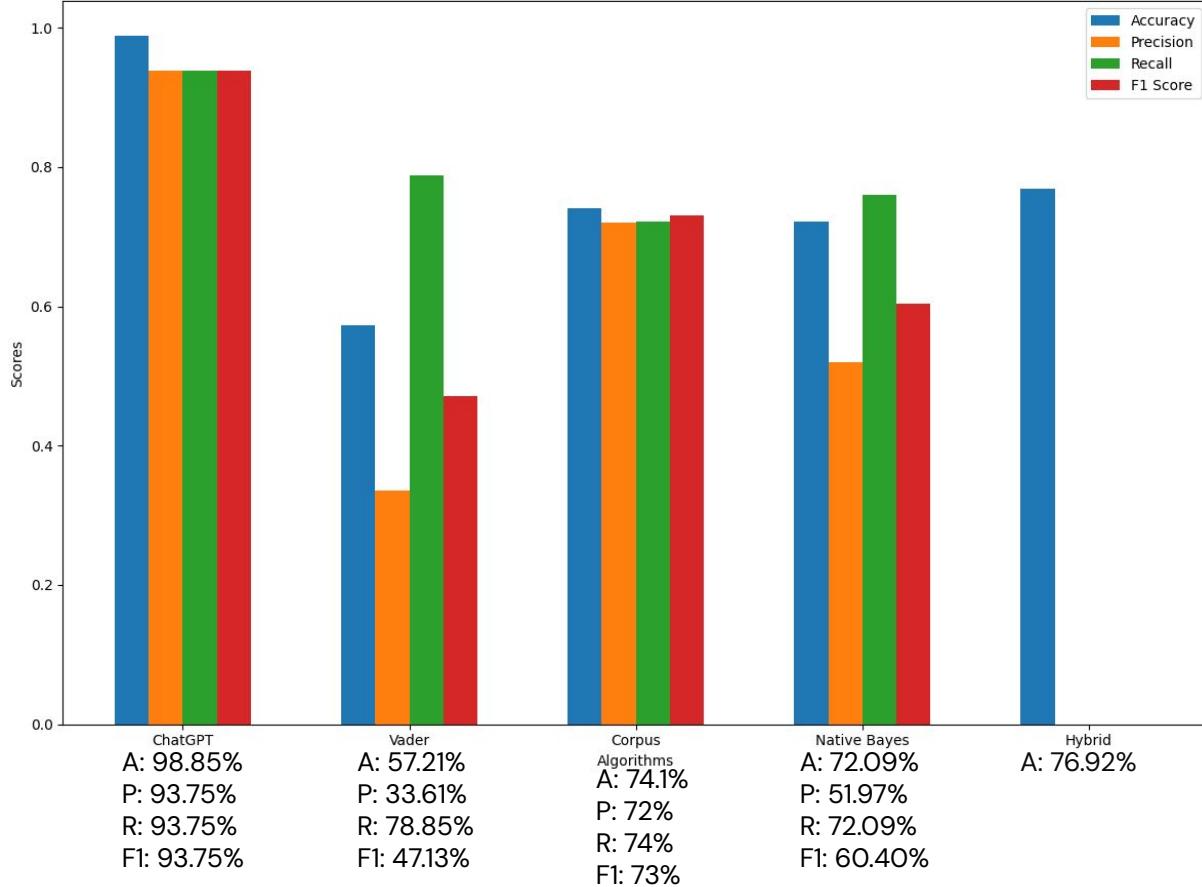
3. Efficiency 

2. Robustness 

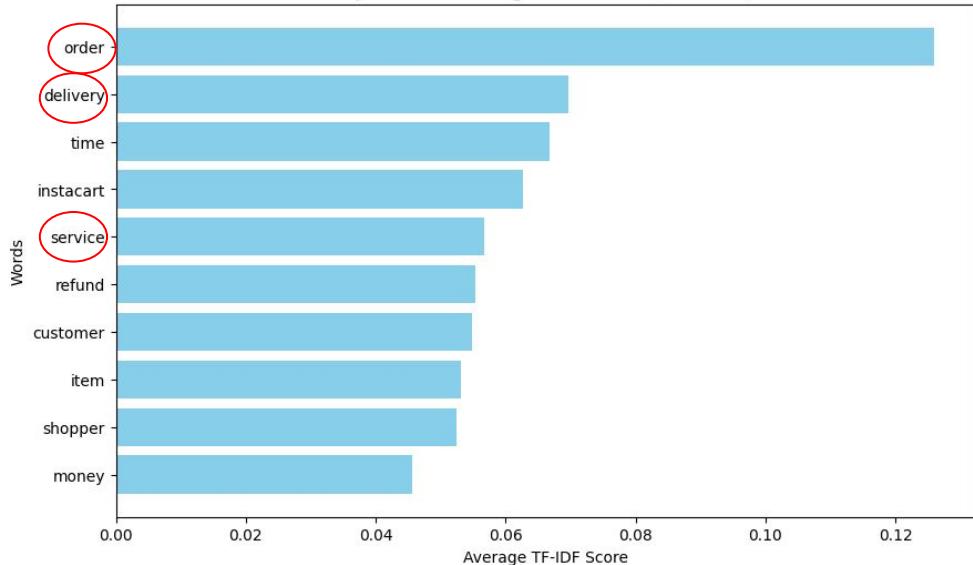
4. Context 

Results

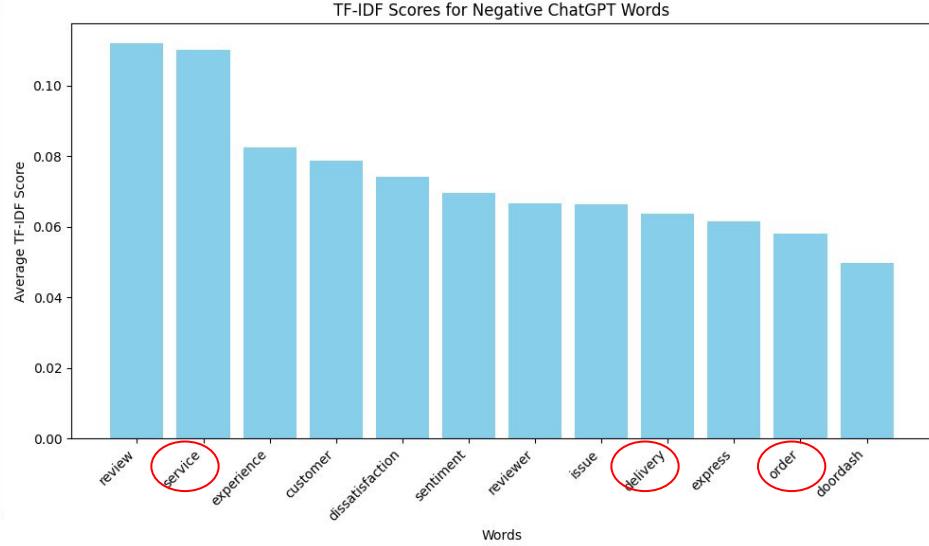
Comparison of Sentiment Analysis Metrics



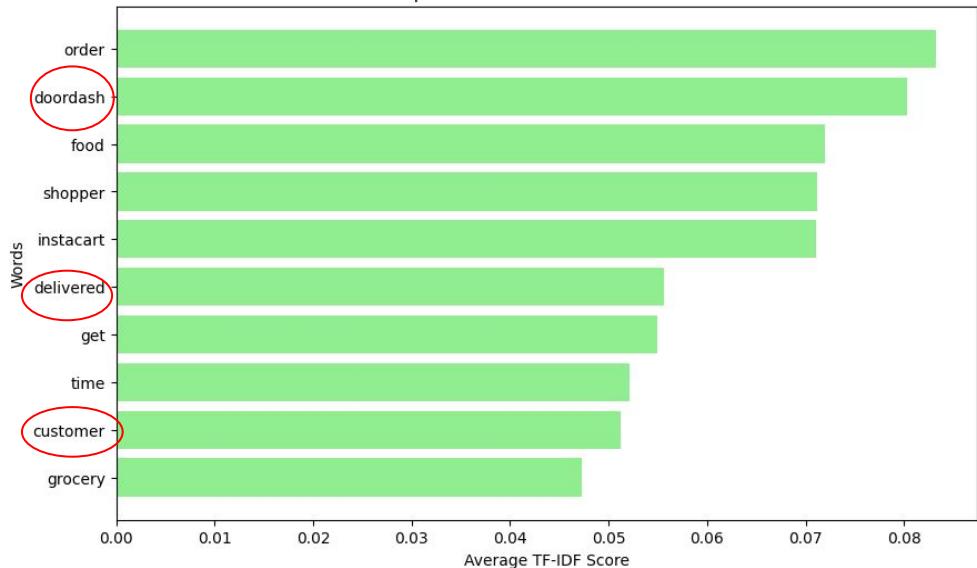
Top Reasons for Negative Reviews (Stars 1-3)



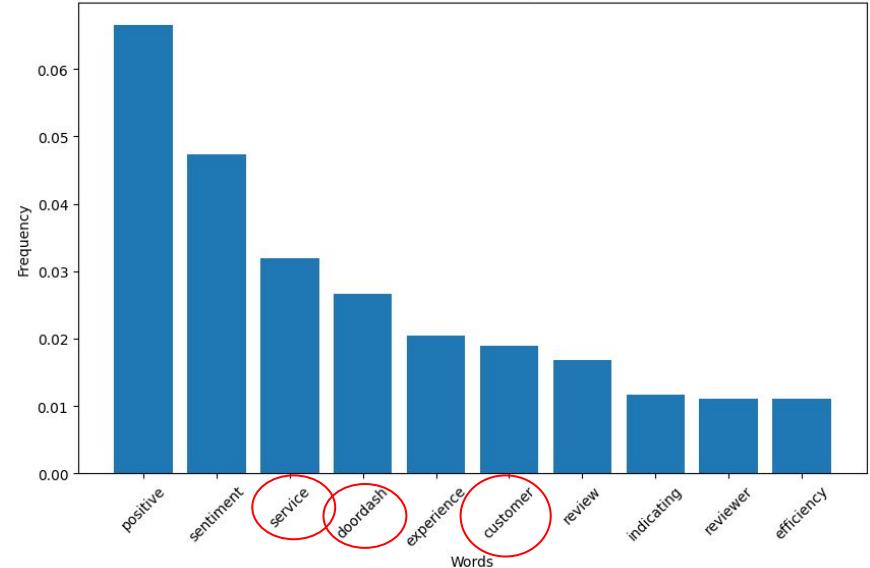
TF-IDF Scores for Negative ChatGPT Words



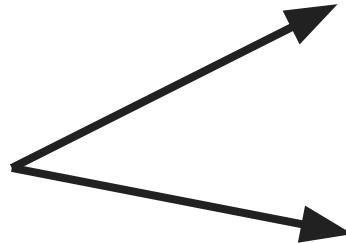
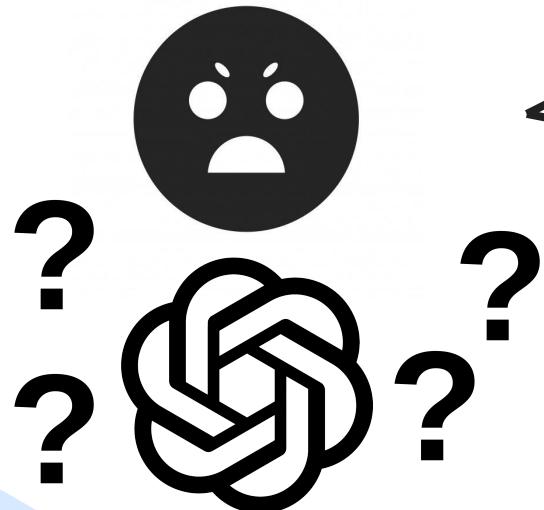
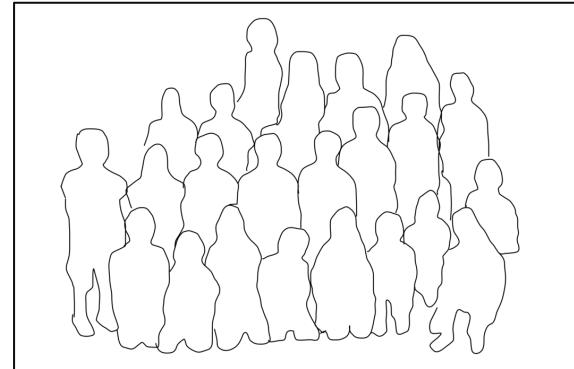
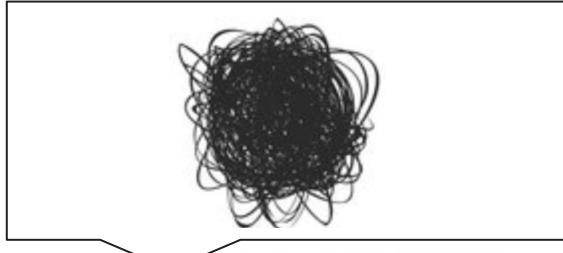
Top Reasons for Positive Reviews



ChatGPT 0 - Word Frequencies



Error Analysis



Model	Interpretability	Efficiency	Robustness	Context
ChatGPT	Low	Low	High	High
Hybrid (VADER & Native Bayes)	Low	Low	High	None
VADER	High	High	Moderate	None
Naive Bayes (NB)	Moderate	High	Moderate	None
Bag-of-Words x Native Bayes	Moderate	High	Low	None

Hypothesis

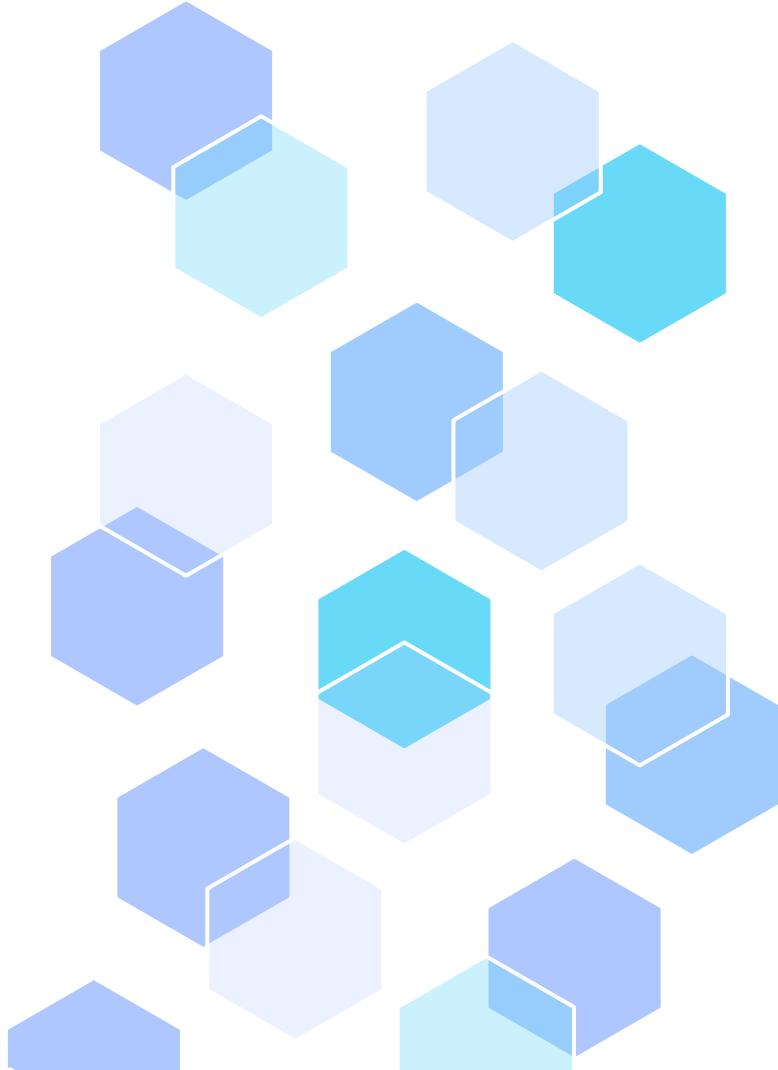
“ChatGPT’s sentiment analysis will outperform
the Top AI algorithms in all metrics.”



“ChatGPT outperforms the Top algorithms in
accuracy, but not efficiency.”

Implications

1. Ethical Concerns
2. New Perspective of ChatGPT
3. Customer & Shopper Satisfaction



Limitations

1. Neutral/Mixed Sentiment Performance
2. Limited Sample Size
3. Manual Analysis



Grocery Apps, ChatGPT, Sentiment Analysis

Ethan Hong



"Biden's Labor Department Takes Aim at Uber and DoorDash with New Rule Aimed to Cut Down on Number of 'independent Contractors'." Fortune, 9 Jan. 2024, fortune.com/2024/01/09/biden-labor-department-rule-independent-contractors-gig-economy-uber-lyft-doordash/.

"Comparing Busines Process, Indicator Performans Metrics and Measuring, Testing Analysis Charts Benchmarking Concept Flat Vector Illustration." Vecteezy, 25 Oct. 2022, www.vecteezy.com/vector-art/13419132-comparing-busines-process-indicator-performans-metrics-and-measuring-testing-analysis-charts-benchmarking-concept-flat-vector-illustration. Flickerbox. "Enhancing the Ratings Experience with Shopper Feedback." Instacart Corporate, 18 Aug. 2021, www.instacart.com/company/shopper-community/enhancing-the-ratings-experience-with-shopper-feedback/.

Freer, Anne. "Food Delivery Apps Have a Customer Service Issue." Business of Apps, 2 Sept. 2020, www.businessofapps.com/news/food-delivery-apps-have-a-customer-service-issue/.

---. "Food Delivery Apps Have a Customer Service Issue." Business of Apps, 2 Sept. 2020, www.businessofapps.com/news/food-delivery-apps-have-a-customer-service-issue/.

Fu, Ziru, et al. "Efficacy of Chatgpt in Cantonese Sentiment Analysis: Comparative Study." Journal of Medical Internet Research, U.S. National Library of Medicine, 30 Jan. 2024, www.ncbi.nlm.nih.gov/pmc/articles/PMC10865189/#:~:text=Our%20findings%20revealed%20ChatGPT's%20remarkable,which%20had%20an%20accuracy%20of.

Griffiths, Ali. "Introducing Updates to Our Ratings System." Instacart Corporate, 15 Nov. 2023, www.instacart.com/company/shopper-community/introducing-updates-to-our-ratings-system/.

Iype, Sujith M. "Human Brain Vs Artificial Intelligence Systems." Ignitarium.com, 5 June 2018, ignitarium.com/human-brain-vs-existing-artificial-intelligence-systems/.

---. "Human Brain Vs Artificial Intelligence Systems." Ignitarium.com, 5 June 2018, ignitarium.com/human-brain-vs-existing-artificial-intelligence-systems/.

JESSICA BURSZTYNSKY. "Fastcompany.com." Fastcompany.com, www.fastcompany.com/90915845/doordash-just-announced-its-biggest-app-update-ever-heres-whats-new.

Lodni, Christin's "The effect of Covid-19 pandemic on the German e-grocery industry with respect to challenges for retailers and customer satisfaction."

Srivastava, Gautam's "A study on reviews of online grocery stores during COVID-19 pandemic using sentiment analysis."

Shaeeli, N. Sakinah's "Customer reviews analytics on food delivery services in social media: a review."

'Tired of Being Stepped On': Instacart Workers Urge Customers to Delete App." The Guardian, 5 Oct. 2021, www.theguardian.com/us-news/2021/oct/05/instacart-shoppers-working-conditions-customers-delete-app#:~:text=As%20Instacart%27s%20workforce%20ballooned%2C%20adding,pay%20rate%2C%20according%20to%20Schaal.

Nast, Condé. "Everyone's Ordering Delivery, but Apps Aren't Making Money." WIRED, 29 May 2020, www.wired.com/story/everyones-ordering-delivery-apps-not-making-money/.