



# Products Advertisement Analysis

Xinrui Wang  
Brown University Data Science Initiative  
Github repository: [https://github.com/xinrui-wang1/products\\_analysis](https://github.com/xinrui-wang1/products_analysis)

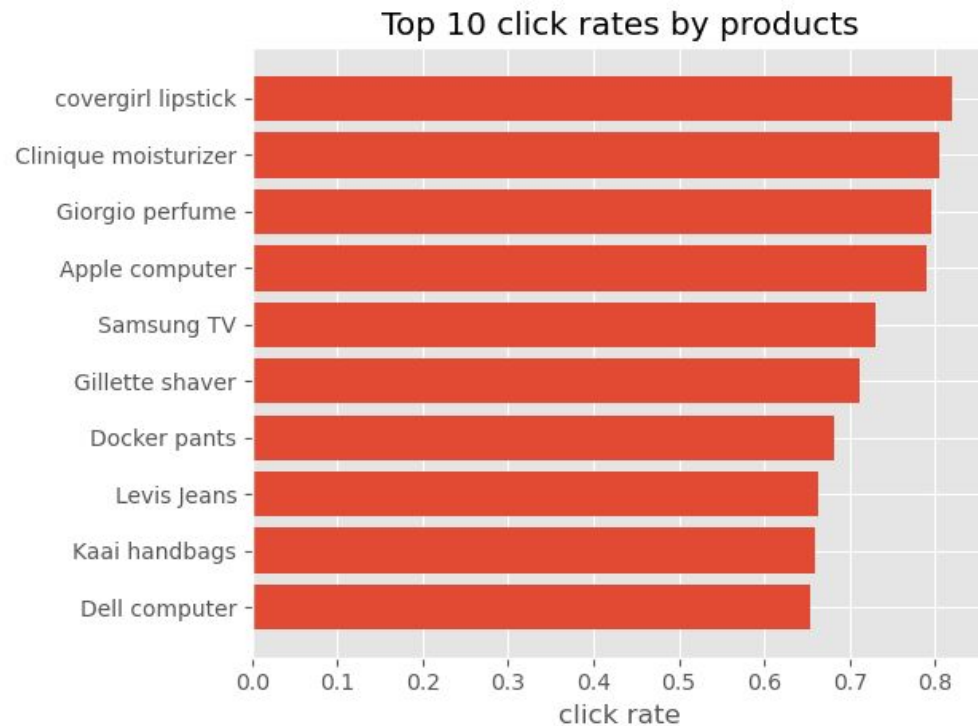


# Data Preprocessing

- Misspelling in URLs
  - Replace with the most similar url (with threshold) based on Jaccard Similarity

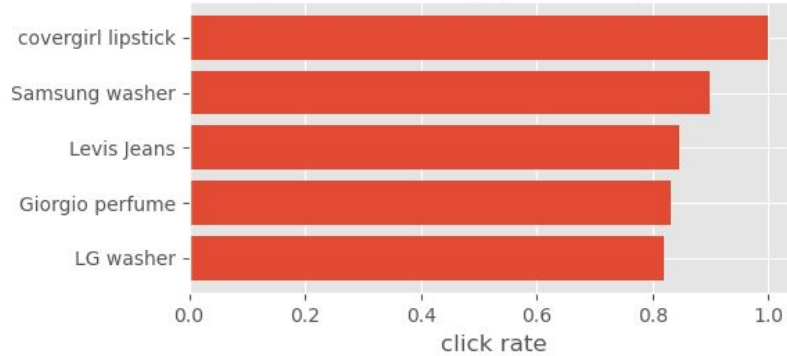
```
Corrupted: http://maybellije.com/lipstick | Cleaned: http://maybelline.com/lipstick  
Corrupted: http://maybelline.com/lipstuck | Cleaned: http://maybelline.com/lipstick  
Corrupted: http://maybelline.com/xipstick | Cleaned: http://maybelline.com/lipstick  
Corrupted: http://nejoK.co/blenders | Cleaned: http://nemoK.co/blenders  
Corrupted: http://nemoK.co/blendess | Cleaned: http://nemoK.co/blenders
```

## Click rate for each product

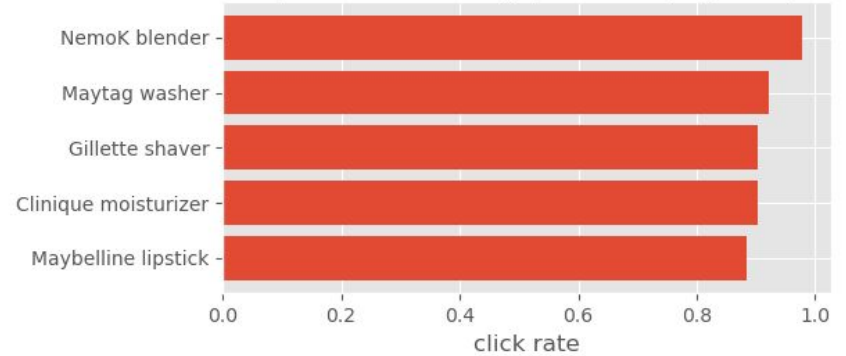


# Top products by click rate/sentiment

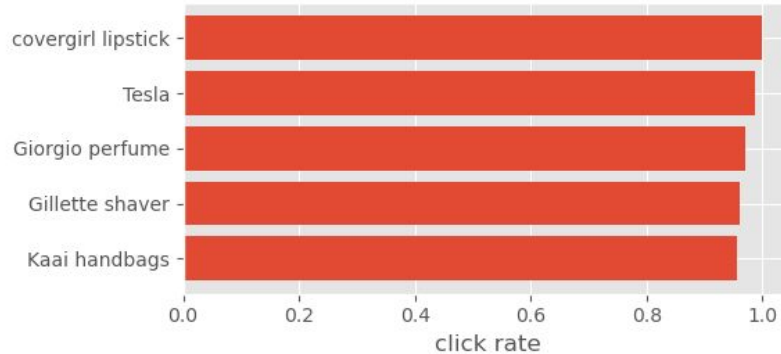
Top 5 click rates by products (positive)



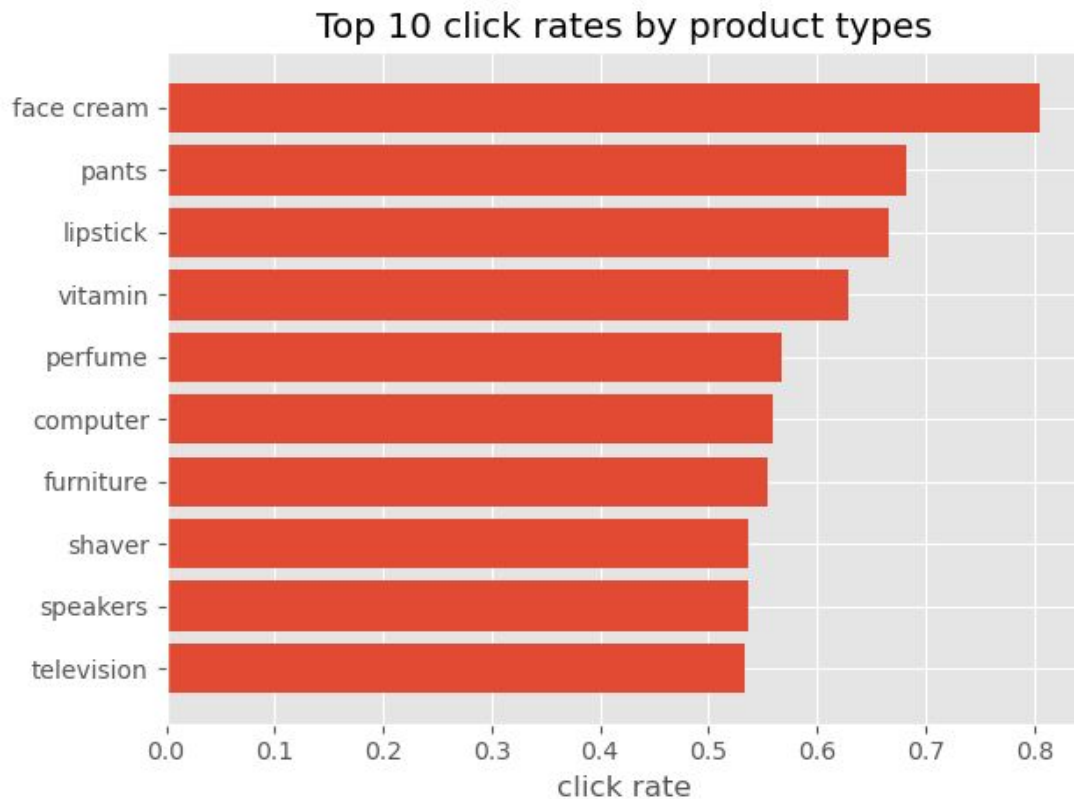
Top 5 click rates by products (negative)



Top 5 click rates by products (neutral)

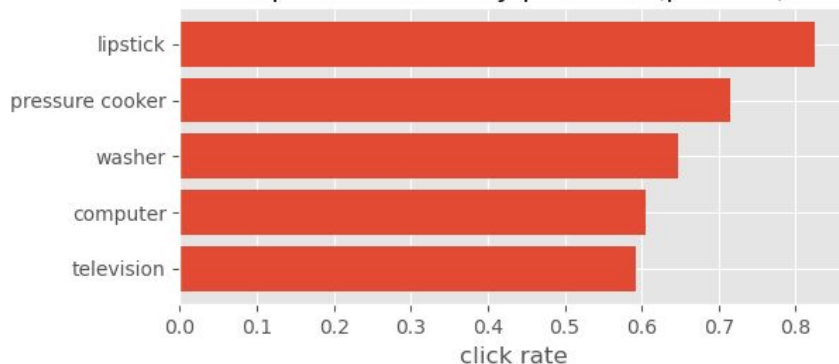


## Click rate for product types

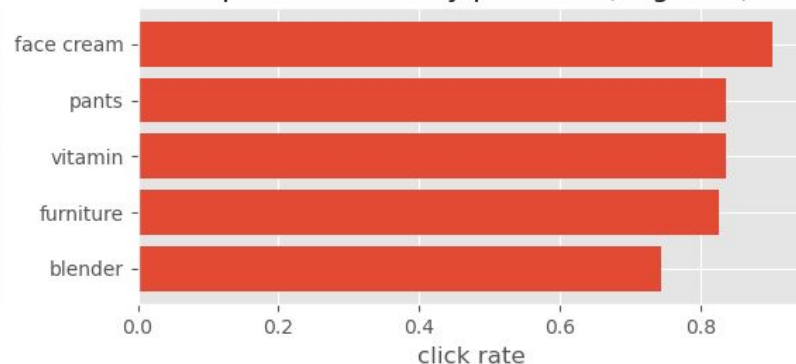


# Click rate by product type/sentiment

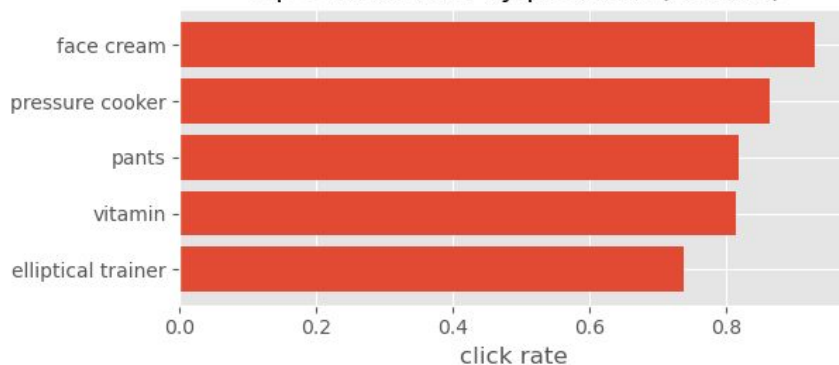
Top 5 click rates by products (positive)



Top 5 click rates by products (negative)

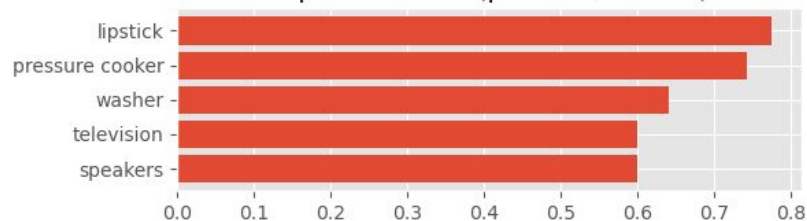


Top 5 click rates by products (neutral)

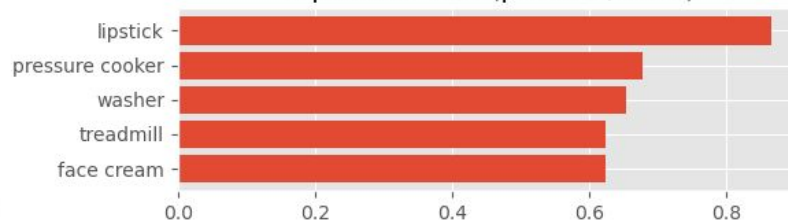


# Top products by gender/sentiment

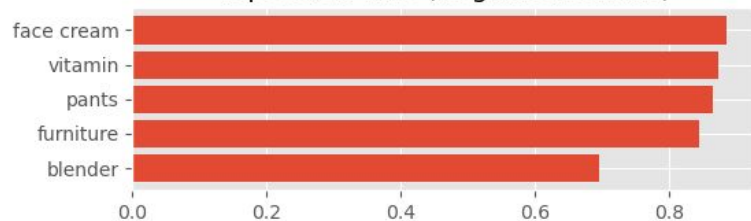
Top5 click rate (positive/female)



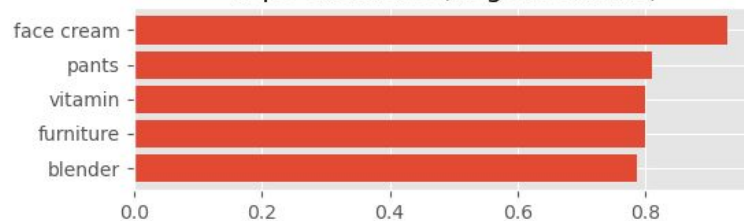
Top5 click rate (positive/male)



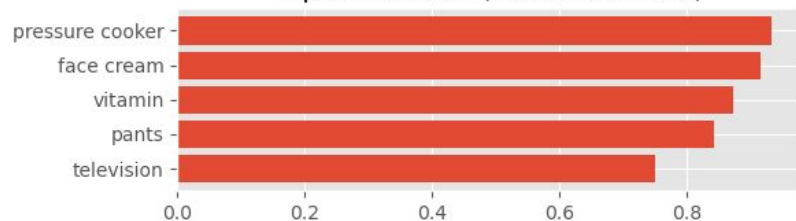
Top5 click rate (negative/female)



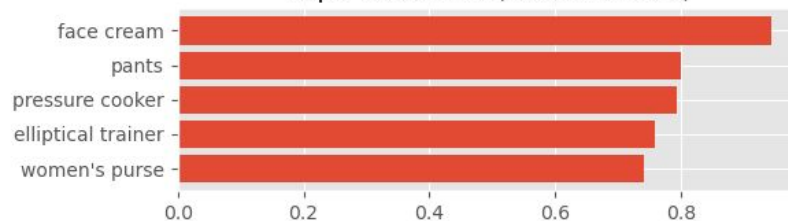
Top5 click rate (negative/male)



Top5 click rate (neutral/female)



Top5 click rate (neutral/male)





## Click rate age groups/sentiment

- Young: face cream + positive sentiment (click rate = 1.00)
- Middle-age: pants + neutral sentiment (click rate = 1.00)
- Senior: face cream + neutral sentiment (click rate = 0.93)
- Juvenile: face cream + negative sentiment (click rate = 0.95)





## Recommendations

- 1) Face cream is one the most clicked products. We should target this product type in the future.
- 2) Sentiment plays an important role in what product is clicked. We should put specific products for different sentiment context.
- 3) Different demographic groups have different click rates for products in different sentiment contexts. Now that we have a look-up table, we should figure out which product is most likely to be clicked per group to best reach our target audiences.



**Thank you!**