



Analysis of Job Advertisements in Social CRM industry

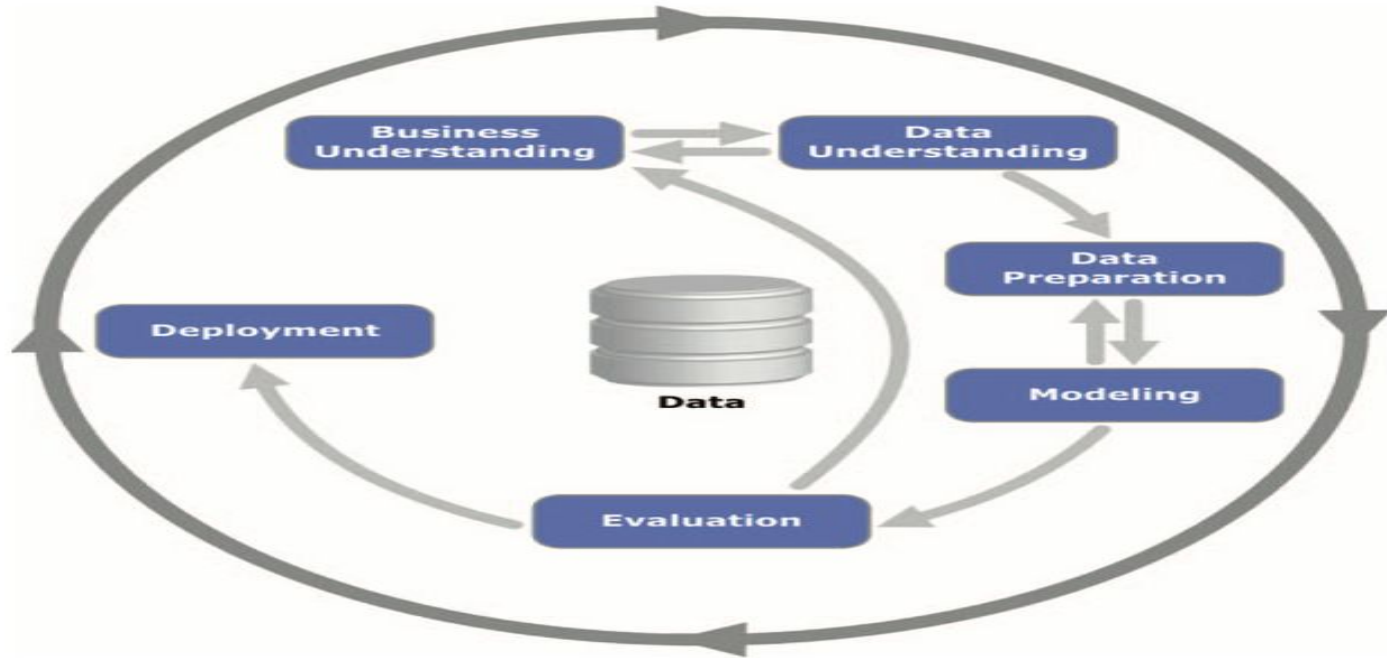
- Jhanvi Lotwala



Problem Definition

- Text Analysis in Job Postings
- Identify knowledge and skills
- Mapping the competencies in the frequently changing job market
- Insight on Social CRM industry
- Helpful in job search and recruitment

Baseline



Dataset variables: job id, job title, job requirements, job summary, job tasks; rows: 1296

Data Cleaning/Preprocessing

- Removal of duplicates based on id and pageUrl
- Conversion of non-English ads
- Removal of common English stop words and some additional stop words like “view”, “e g”, “please”, “send”, “end”, “de”, etc.



Data Extraction

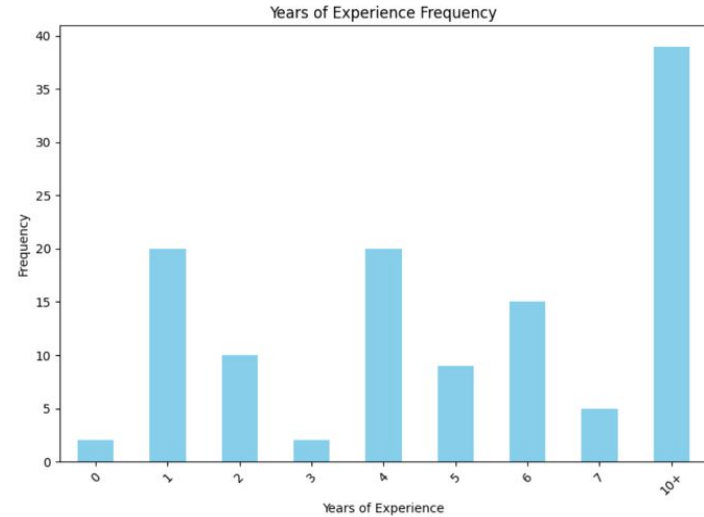
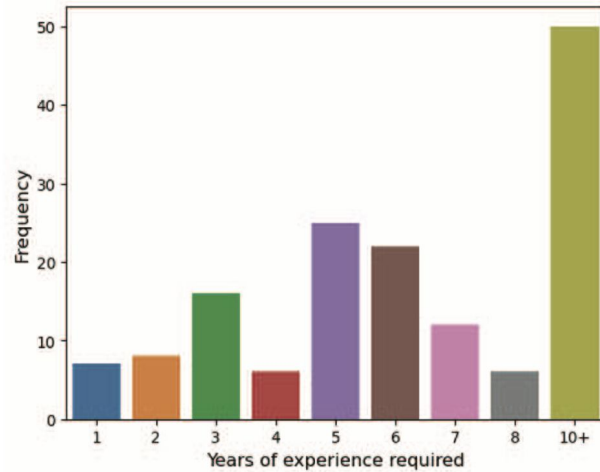


Fig. 4. Years of experience required.

Word Cloud

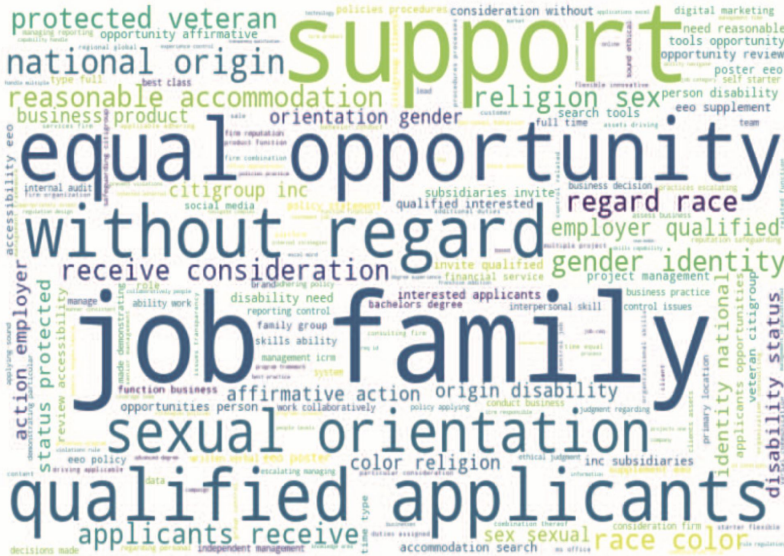
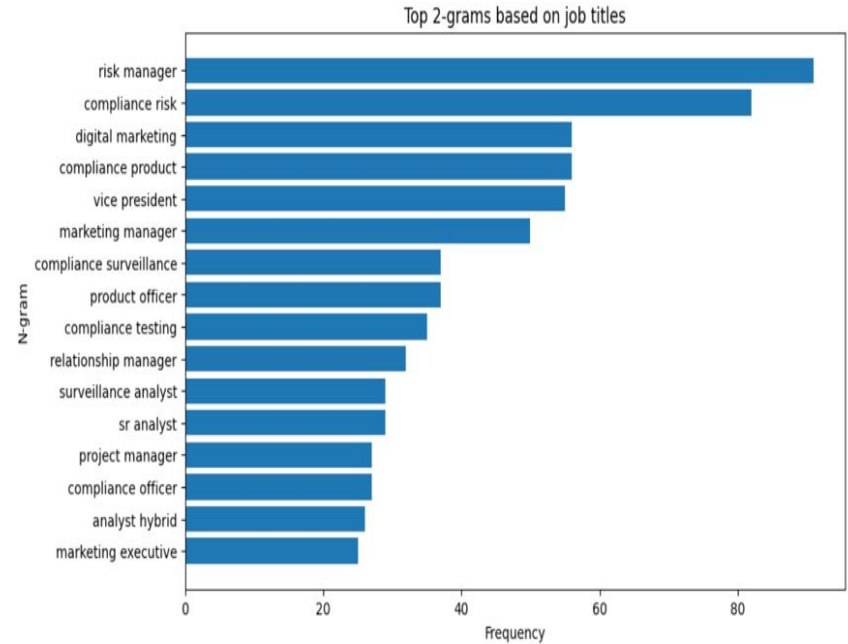


Fig. 3. Word cloud representation of job posts descriptions.

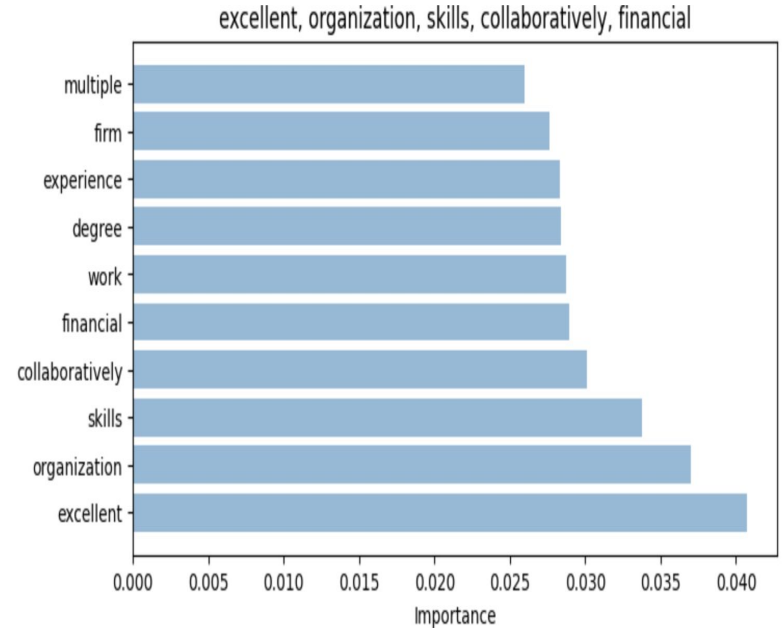
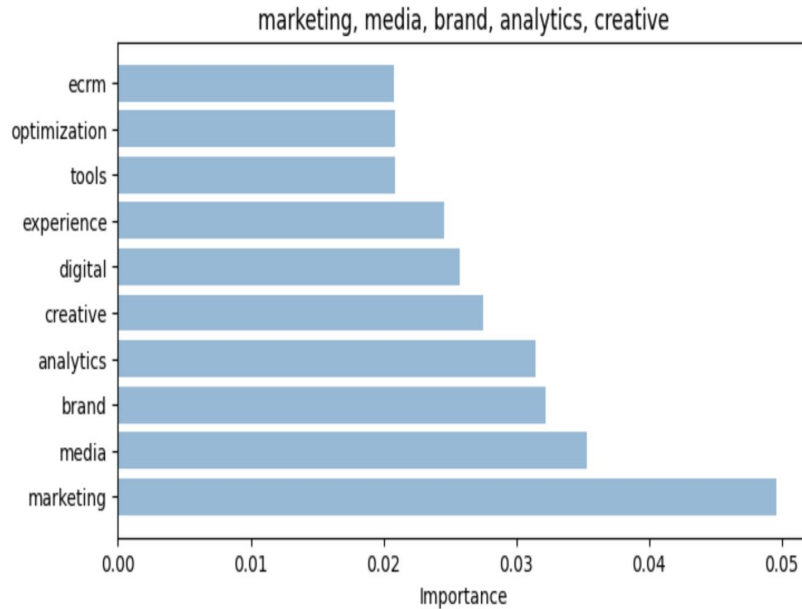
N-grams



Fig. 5. Representation of N-grams constructed based on job titles.



BERTopic Modeling



BERTopic Modeling

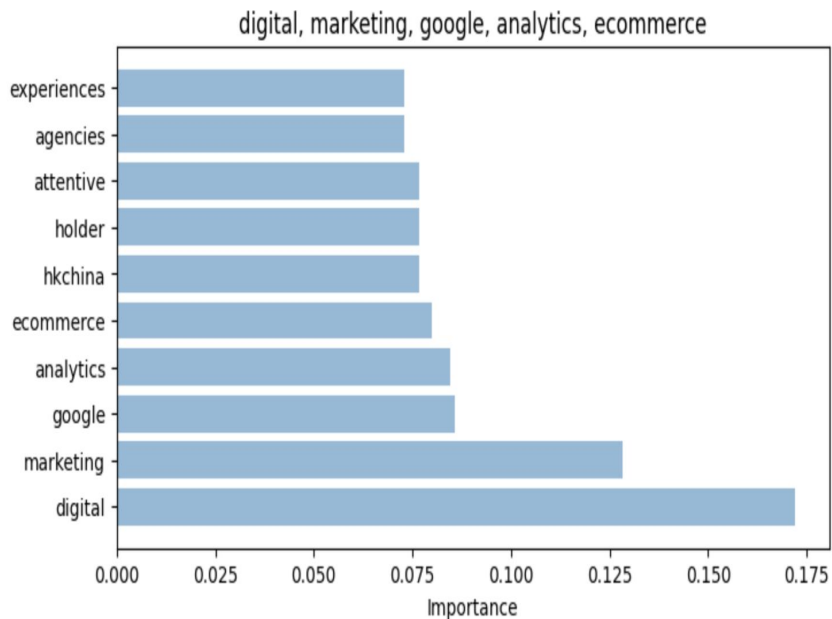
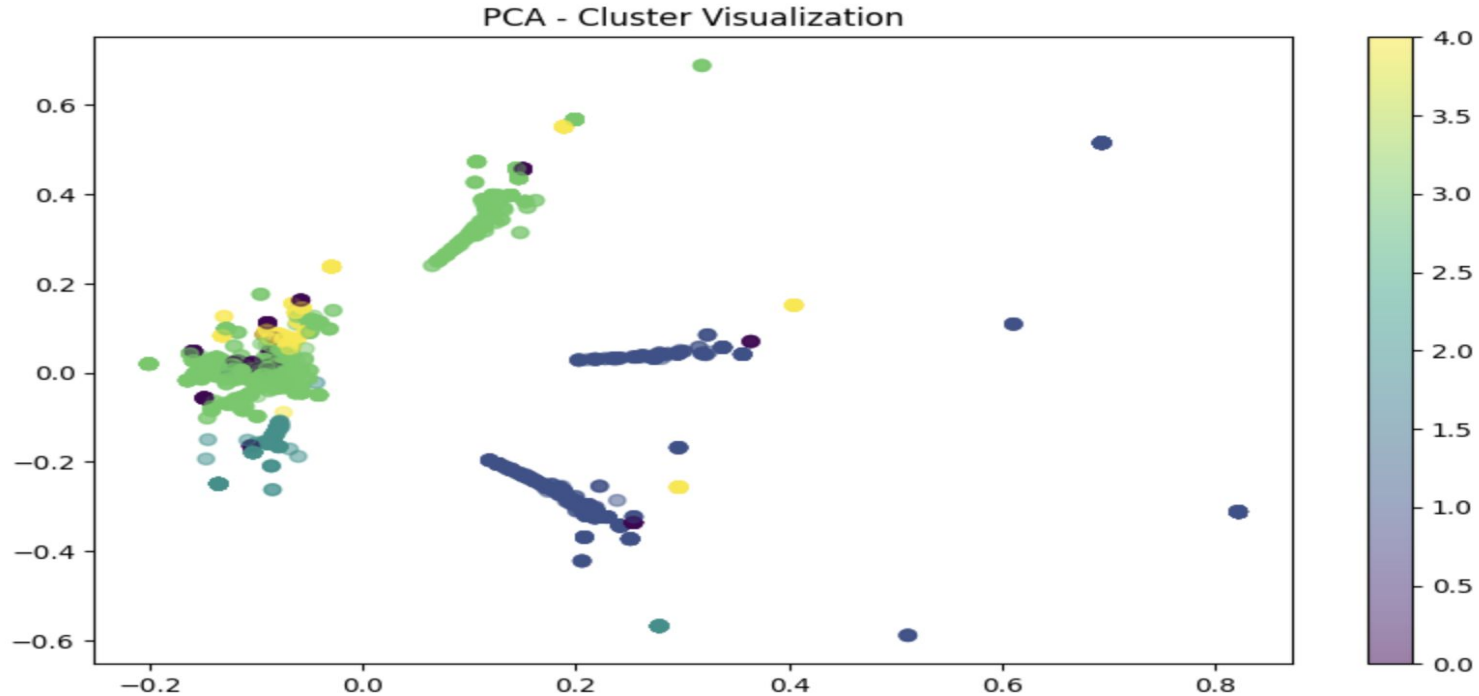
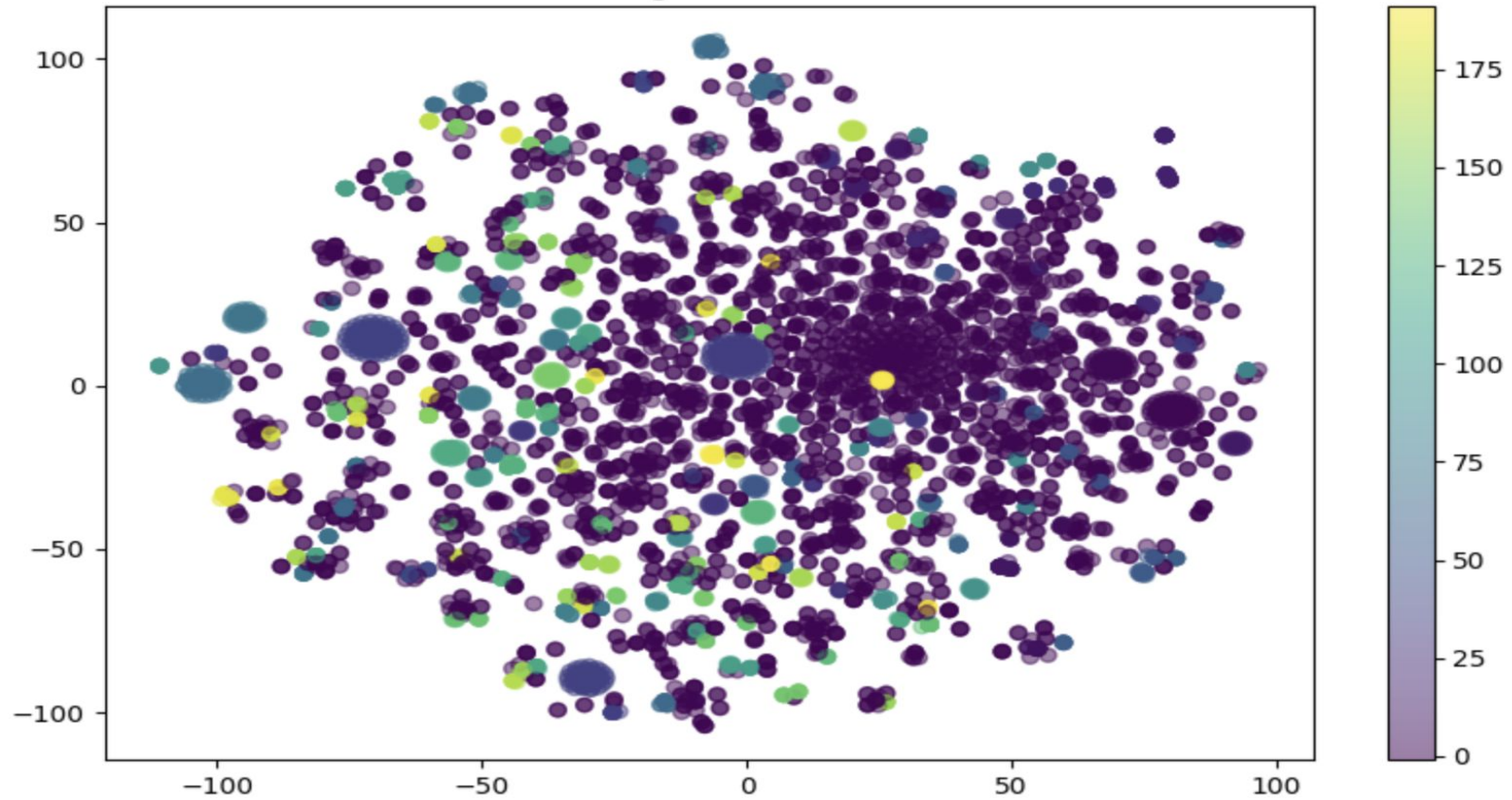


Fig. 9. Main knowledge required

Clustering w/ Dimensionality reduction



DBSCAN Clustering with t-SNE visualization



Evaluation/Discrepancies

- N-grams, Word Cloud, and, Years of Experience showed good results
- Evaluated each text column
- Baseline Approach: BERTopic Modeling
- New Approach: Clustering along with dimensionality reduction techniques
- Scraped data had some rows with neither English nor any other language
- Some noise reduction needs to be done manually

References

[1] P. C. Menezes, A. F. Justino, B. A. Barata, A. F. Jacob Junior and F. M. Lobato, "Market Overview in Social CRM: An Analysis of Job Advertisements," in *2023 IEEE/WIC International Conference on Web Intelligence and Intelligent Agent Technology (WI- IAT)*, Venice, Italy, 2023, pp. 555-562, doi: 10.1109/WI-IAT59888.2023.00092. keywords: {Industries; Social networking (online); Taxonomy; Customer relationship management; Companies; Market research; Intelligent agents; Data Analysis; Social CRM; Job Analysis; Text Mining; Market Analysis}

[2] A. Mueller, "wordcloud," PyPI, Jul. 26, 2018. <https://pypi.org/project/wordcloud/>

[3] M. P. Grootendorst, "Home," *maartengr.github.io*, 2022. <https://maartengr.github.io/BERTopic/index.html#quick-start>



Thank You for listening:)