Fintech Hiring trends in the largest banks in the US

Written by Sri Krishnamurthy, CFA

Financial institutions in the US are changing rapidly. Changing business models and the technological revolution has fueled the growth of a new breed of financial products and services collectively known as Fintech. With changing demographics, automation efforts and demand for new products and services, large financial institutions are realizing the power of technologies like data science, AI, cloud technologies and machine learning and are heavily investing to upgrade their technological platforms to cater to the upcoming revolution. Technology has been a key player in helping drive this revolution. In a 2017 report by CB Insights (See appendix) more than 46% of the job openings were in the technology section. Things are fast evolving and as we enter 2019, it is interesting to understand the hiring trends in the top financial institutions in the US.

Our goal in this case study is to conduct a study on the job openings in the top US Banks in the United States and analyze trends in the industry particularly in the area of Fintech.

Methodology:

1. We will create three csv lists of top 100 keywords we can use to describe fintech. Each list will be of the format

RowId	Word
1	Python
2	Fintech

2. We will then scrape job openings from the top 24 banks in the US and look for occurrences of the top 100 keywords in each job posting. We will then build a dataset that will look something like this.

Job no	Institution	List id (1,2,3)	URL (URL Of job posting)	List id (1,2,3)	1	2	.....	99	100
	J P Morgan	1	https://jobs.jpmorganchase.com/ShowJob/Id/191925/Python-Software-Engineer-Digital-Investment-Bank/	1	0	5		1	4

3. We will collect additional data as needed then analyze this dataset to understand hiring trends how the emphasis the institution has on key fintech areas.

PART 1: (30)

Understanding keywords that are frequently used in Fintech

Our first objective is to understand the keywords that are typically used in Fintech. The World Economic forum (WEF) has put together research in the area and it is summarized in 4 reports.

1. http://www3.weforum.org/docs/Beyond_Fintech_-_A_Pragmatic_Assessment_of_Disruptive_Potential_in_Financial_Services.pdf
2. http://www3.weforum.org/docs/WEF_The_future__of_financial_services.pdf
3. http://www3.weforum.org/docs/WEF_The_future_of_financial_infrastructure.pdf
4. http://www3.weforum.org/docs/WEF_A_Blueprint_for_Digital_Identity.pdf

These four reports comprehensively discuss the state of the industry and the upcoming trends.

We will extract keywords from these documents to build a dictionary of top keywords that could be used to describe the fintech space. This needs to be done in 2 steps.

1. Convert the report to text, clean it and store it in a csv file
   You can use packages like Pypy2 to do it. See https://www.geeksforgeeks.org/working-with-pdf-files-in-python/ for examples
2. We will extract the top 100 keywords and build three lists using the following approaches:
   a. Wordcount (count all words and rank by frequency and choose the top 100)
   b. TF/IDF
   c. TextRank

   Note:
   - Remove special characters
   - Remove stopwords using the long list from https://www.ranks.nl/stopwords
   - Manually review the top 100 words and remove names of people, companies, words that don't make sense etc. and curate the three lists to usable lists.

   See some examples that may help with your implementation here: (Note: I just googled. Can't guarantee the code would work)
   - https://github.com/JRC1995/TextRank-Keyword-Extraction
   - https://summanlp.github.io/textrank/
   - https://xang1234.github.io/textrank/

Deliverables:

- Three csv files of top 100 key words for each methodology (2a, 2b, 2c)

PART 2: (30)

The list of the top US Banks and their summary stats are available here: https://www.relbanks.com/top-us-banks/market-cap. Your team with the assigned team number will analyze data from 2 banks with the algorithm (team number*2-1, team number*2) (Ex: If team 1, do bank 1,2. If Team 12, do bank 23,24).

1. For each of the two companies, build a scraper to scrape the data from the respective company's careers site.
- Note that you may have to build a scraper from scratch. Popular tools are BeautifulSoup, Scrapy etc. You can see https://github.com/amrrs/Scraper-Projects for some examples.
- See https://github.com/karenhao/techreview_arxiv_scrape for inspiration

2. Parse each job posting and do keyword extraction after removing special characters and stem words. Each posting should have a dictionary of words and the frequency of occurrences.
3. Collect any other company, position, location specific information in datasets (tables/json/csv files)ß that may be needed for analysis.
4. Using each of the three lists created in Part 1, build a dataset which has this format

Job no	Institution	List id (1,2,3)	URL (URL Of job posting)	List id (1,2,3)	1	2	.....	99	100
	J P Morgan	1	https://jobs.jpmorganchase.com/ShowJob/Id/191925/Python-Software-Engineer-Digital-Investment-Bank/	1	0	5		1	4

Deliverable:

- A csv file for the table mentioned above.
- Any other data files you summarized that may be needed for analysis

PART 3: (40)

We will now analyze this data and draw insights on hiring patterns and trends. You are free to calculate any other metrics you may think useful. See Appendix 1 for example. Our final report is a comprehensive description covering these aspects:

1. Key areas of fintech that were inferred from the lists you created in Part 1.
2. Key hiring trends drawn from Part 1 and 2.
3. Your interpretation on actionable insights drawn from your analysis.

You should augment your analysis with graphs and charts. You can use Matplotlib, plotly, seaborn. I would recommend you try https://www.datawrapper.de

Deliverable:

Your final report will be in the form of a claat document (See https://github.com/googlecodelabs/tools for information on how to create claat documents.)

See https://www.technologyreview.com/s/612768/we-analyzed-16625-papers-to-figure-out-where-ai-is-headed-next/ for inspiration.

Note:

- All work referencing other people's code should be cited completely.
- Email analyticsneu@gmail.com and the TAs your github link.
- Your github link should have all the functional code and report
- Datasets you use for your report should be in the **Data** folder on github
- Delivery deadline: 8th Feb 11.59pm

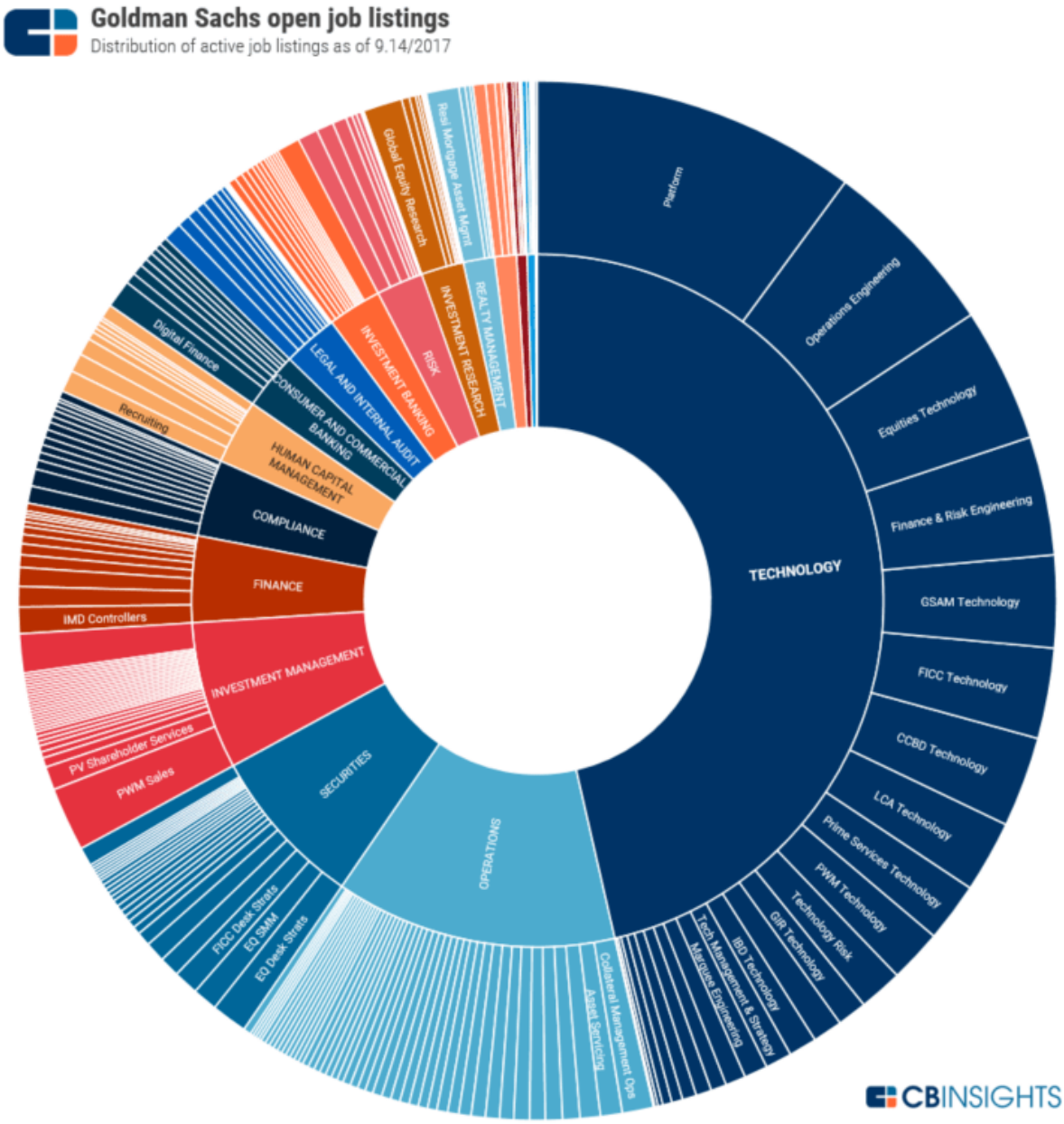
Additional clarification – Jan 29 2019

- Please see sections in yellow

Appendix 1:
Source: https://www.cbinsights.com/research/report/goldman-sachs-strategy-teardown/

Hiring trends

An analysis of Goldman's 2,000+ open job listings by division and business unit confirm its focus on technology and digital finance. As of 9/14/2017, 46% of all of Goldman's jobs were in its technology division. The highest percentage of technology jobs were for platform roles, followed by operations engineering and equities technology positions. 6% of all technology jobs at Goldman Sachs were in its Consumer and Commercial Banking business unit, which houses its Digital Finance subdivision.



Some key takeaways from analyzing Goldman's hiring data:

- **Goldman is actively hiring Android developers** to enable the creation of its "all digital retail bank." Earlier job postings focused on hiring iOS developers. This hints that Goldman's Digital Finance (GS Bank, Marcus) product offerings may soon include native mobile apps.
- **Goldman is hiring for roles to build out its own robo-advisory platform**. A job posting for a user experience engineering position notes that Goldman's "Digital Advise Solutions (DAS) business cover mass affluent market by building an Automated Digital Advice Platform."
- **Goldman is bulking up hiring for its Digital Finance unit in the UK**. London-based roles in the company's Digital Finance subdivision include communications, customer support, product management, and DevOps.
- **Goldman counts zero public-facing job postings mentioning the term "blockchain"** at present, despite setting up a microsite dedicated to the technology titled "Blockchain: The New Technology of Trust."
- **Goldman is also making a number of engineering hires for its Marquee platform**, which provides clients access to its analytics, trading, and data tools. 15+ engineering jobs in New York or India are for developer roles related to Marquee engineering. The hiring comes as Goldman is now looking to monetize some of its software as well. Goldman is also reportedly seeking investments to spin out its web app Simon, which allows brokers to buy and adjust structured note products, at a valuation of around $75M.