**Assignment: -**

1. **A developer is assigned a task to scrape 1 lakh website pages from a directory site, while scrapping he is facing such hcaptcha, which are placed to stop people from scrapping As a project Coordinator suggest ways to solve this problem**

**Ans:**

1. [Use a CAPTCHA solving service such as 2Captcha or Anticaptcha](https://ieeexplore.ieee.org/document/9001774/" \t "https://www.bing.com/_blank).
2. [Use computer vision techniques such as OpenCV to process the CAPTCHA images and extract the text.](https://medium.com/@ageitgey/how-to-break-a-captcha-system-in-15-minutes-with-machine-learning-dbebb035a710" \t "https://www.bing.com/_blank)
3. [Use deep learning frameworks such as Keras and TensorFlow to train a neural network to recognize the characters in the CAPTCHA images](https://medium.com/@ageitgey/how-to-break-a-captcha-system-in-15-minutes-with-machine-learning-dbebb035a710" \t "https://www.bing.com/_blank).
4. [Use Python to scrape LinkedIn profiles and extract data such as job titles, degrees, and majors to predict salaries using linear regression or random forest models.](https://cheq.ai/blog/how-fraudsters-bypass-captchas/" \t "https://www.bing.com/_blank)
5. [Use LinkedIn Talent Insights to view talent pool reports that provide compensation information inferred using LinkedIn Salary data collected from LinkedIn members and estimated based on similar roles, companies, and regions.](https://github.com/topics/hcaptcha-bypass" \t "https://www.bing.com/_blank)
6. **Our client has around 10k linked-in people profiles, he wants to know the estimated income range of these profiles. Suggest ways on how to do this?**

**Ans:**

1.[Use the LinkedIn Salary Insights feature to get compensation information based on similar roles, companies, and regions 1](https://www.linkedin.com/help/talent-insights/answer/a186045/salary-insights-in-talent-pool-reports?lang=en" \t "https://www.bing.com/_blank).

2. Use data science techniques such as machine learning or data analysis to estimate income ranges. [For example, you can use statistical modeling to estimate the quantiles (10th and 90th percentiles, and median) and histograms for base salary, bonus, and other compensation types 2](https://engineering.linkedin.com/blog/2017/12/statistical-modeling-for-linkedin-salary" \t "https://www.bing.com/_blank). [You can also use Python to scrape LinkedIn profiles and extract data such as job titles, degrees, and majors to predict salaries using linear regression or random forest models 3](https://github.com/sindhri/salary_prediction" \t "https://www.bing.com/_blank)[4](https://www.scrapingdog.com/blog/scrape-linkedin-profiles-using-python/" \t "https://www.bing.com/_blank).

3. [Use LinkedIn Talent Insights to view talent pool reports that provide compensation information inferred using LinkedIn Salary data collected from LinkedIn members and estimated based on similar roles, companies, and regions 1](https://www.linkedin.com/help/talent-insights/answer/a186045/salary-insights-in-talent-pool-reports?lang=en" \t "https://www.bing.com/_blank).

4. [Use LinkedIn’s new salary estimating feature, which was built based on data from their members and employer information, to see an estimated or expected salary.](https://www.listfoundation.org/how-to-find-the-salary-range-for-a-linkedin-job-listing" \t "https://www.bing.com/_blank)

1. **We have a list of 1L company names, need to find linked-in company links of these profiles, how to go about this?**

**Ans**:

Here are the steps to use the LinkedIn Company URL Finder tool:

1. Create a free PhantomBuster account.
2. Give the names of the companies whose LinkedIn pages you want to find.
3. Specify the number of names to process per launch.
4. Set the Phantom on repeat.
5. Download a CSV spreadsheet or JSON file of all your LinkedIn company pages.
6. **How to identify list of companies whose tech stack is built on Python. Give names of 5 companies if possible, by your suggested approach.**

**Ans:**

1. Search for “companies whose tech stack is built on Python” on a search engine.
2. Look for articles or blog posts that list companies using Python in their technology stack.
3. Check the credibility of the source and verify the information provided.
4. Compile a list of companies that use Python in their technology stack.
5. **Need to find an API, through which we can send linked-in messages to other linked-in users**

**Ans:**

1. Create a Linked-in account and a Linked-in Developer account.

2. Create an application on the Linked-In Developer Portal.

3. Obtain an access token by following the authentication process.

4. Use the access token to authenticate your requests to the Linked-in Messaging API.

5. Send messages to other Linked-In members using the API.