Capstone Project Batch-15 Weekly Report

Date: 05/06/2022

Project Details:

1. Sponsor Company: Peak Health Technologies

2. Project Title: Using a combination of multiple facets of Health, Engagement and Productivity to estimate occupational burnout

Project Milestones:

Progress made in Current Week:

- Exploratory Data Analysis
  - Explored review text data
  - Data cleaning to extract Location and Job Experience
- Pre-Processing of text data and creation of word clouds
- Sentiment Analysis using Textblob
- First mock-up draft of State of the Workforce dashboard

Tasks to finish in Next week:

- Mapping given metrics and data onto three major axes of burnout
- Evaluating scores for five different personas based on the burnout dimensions
- Explore novel ways of visualizations on the dashboard

Updates/MoM from Sponsor and Faculty Mentor:

Note: It is expected that you have at least one weekly connect with the faculty mentor and sponsor. If you were not able to schedule meetings with the sponsor or faculty mentor in the current week, please mention the reason for your inability to meet with the Sponsor or Faculty Mentor.

Updates/MoM from Sponsor:

- Focus on the numeric variables to finalise the framework and incorporate text data at a later stage
- Map given metrics and data onto three major axes of burnout
- Evaluate scores for five different personas based on the burnout dimensions
- Explore novel ways of visualizations on the dashboard
- Schedule a working session with Tarun to review dimension mapping
- Extract more review data from any public review site.

Updates/MoM from Faculty Mentor:

Meeting scheduled on 06th June, 2022

Challenges:

Mention any technical and non-technical challenges that you faced during the current week that hindered your project progress. Enter "NA" if you didn't face any challenges.

Technical Challenges:

- Identifying reliable sources to extract review data

Non-Technical Challenges:

NA

Mention any other queries/challenges regarding the project that you want to highlight:

Defining the five personas based on the available scores given in the data