SENTIMENT ANALYSIS OF IT COMPANIES- FEATURES AND FEELINGS PROFILING, AS WELL AS DATA MINING FOR EMPLOYEE INPUT

Amaresh Bose, Department of BCA, St. Xavier's College, Dumka
Naghma Khatoon, Usha Martin, University, Ranchi
amaresh10999@gmail.com; naghma.bit@gmail.com

Abstract:
With so many businesses and organizations to choose from, making a final decision may be time-consuming and laborious. There are several metrics available for evaluating companies, but what's required is a metric that takes into account all of the aspects that contribute to employees' overall opinions of the company. To tackle this problem, our study concentrates on creating an aspect-sentiment-based analysis for corporations. To do this, we will rely on the honest feedback of our employees on these companies. We collected a large dataset of corporate reviews from the omnipresent 'Glassdoor.com' and used a novel ensemble approach to analyze the reviews' sentiment across a wide range of variables. In addition, we provide various insights from the dataset to aid customers in better comprehending their options and making educated choices when picking businesses.

Keywords: *Sentiment analysis, Employee opinion, Mining, Information Technology, Dataset.*

1. Introduction:

Interpersonal communication relies heavily on expressing and interpreting a wide range of emotional states, mental dispositions, and value judgments, from the most positive to the most negative. Improving the human-machine interface requires a deeper understanding of human emotions, emotional reasoning, and the linguistic expression of emotions. With the advent of the internet and the ease with which information can be shared, new thought production and dissemination methods have become available in today's age of continual communication. Academics may be interested in the public's perspective on various problems, causes, and events by keeping tabs on the conversation on social media. This is primarily due to two factors: first, exposure to other viewpoints may increase one's decision-making skills. Second, companies may use this data to see how customers feel about their products and services to improve or introduce new ones. Sentiment analysis studies emotions and viewpoints from various sources, including text, images, and audio.[1-3] Up-to-date sentiment analysis techniques make excellent use of linguistic patterns and deep neural networks. The ability to recognize elements inside text is just as important. Aspects are parts of the product that get positive reviews, such as the battery in the line "the battery lasts a long time." The most challenging part of sentiment analysis is identifying unique characteristics and emotions. Our method involves incorporating an ELM classifier with linguistic patterns.

Fig 1: Source (https://hp-analytics.medium.com)

Employees are an organization's most precious asset, and their ideas are crucial to the business's growth. Review sites and job search engines have been inundated with employee suggestions recently. A company must have a system to frequently evaluate employee performance since it creates a more pleasant work atmosphere, boosts morale, and increases employee loyalty. Job-

seekers may use these evaluations to choose organizations that best meet their needs. Despite their importance as data sources for sentiment mining, these assessments have received little attention from the scientific community. As far as we know, this study only included businesses that used Glassdoor ratings.[4] But they could only glean overarching sentiments and themes from the reports. To conduct this study, we compiled a large dataset from Glassdoor, including reviews from Singaporean workers about their employers. Combining ELM with emotional patterns in this dataset allowed for several investigations, including aspect extraction and aspect-based sentiment analysis. We did this by developing company-specific representational embeddings, weighted by the emotional score of its constituent parts. A 24 -dimensional space is used to map out how customers feel about various aspects of a firm. Factors like company culture, salary, and geographic proximity are all examples.

The dataset we used for this research is comprised of 40,000 Glassdoor ratings. Naturally, there will be a broad range of praise and criticism from various reviews, forming an image of the company. The two most significant results of this research are as follows:

- We are putting together an extensive Glassdoor data set for sentiment analysis at the facet level. This dataset results from feedback from current and previous employees of the associated firms.
- They discover patterns among companies from various sectors using firm-specific aspect-sentiment embeddings. The workers' shared perspective creates a virtual recreation of the company's physical space in three dimensions. A company selection is complicated by various factors, including pay, benefits, and schedule flexibility. Job-seekers will find this feature very useful in their search for companies that suit their individual needs.

Features	Features Phase
Vacation	Decent, Mandatory, Paid, Planned, Considering, Balance
Work Life	Competent, Mundane, Exciting, Friendly, Versatile, Stressful
Working time	Exciting, Peak, Tough, Stressful, Extra, Irregular
Job training	Prepares, Competent, Notch, Outstanding, Outdated, Tough
Company support	Excellent, Tedious, Benefits, Supervisors, On site, Rare
Technology	Excellent, Ancient, Global, Latest, Green, Innovative
Official Staff	Understanding, Excellent, Competent, Mean, Motivated, Dysfunctional
Stress	Underpaid, Excessive, Good, Constant, Additional, Incompetent
Job respect	Professional, Mutual, Utmost, Solid, Diminishing, Great
Salary	Optimal, Advancement, Hikes, Midrange, Fantastic, Unattractive
Politics	Dysfunctional, Drive, Dirty, Extreme, Everywhere, Internal
Work projects	Numerous, Diverse, Challenging, Pushed, Creative, Unbearable
Perks/Benefits	Unique, Scares, Incredible, Illusion, Lousy, Incentives
Performance	Personal, Necessary, Measurable, Extraordinary,

	Encouraged, Technical
Market viability	Changing, Shrinking, Impact, Competitive, Unknown, Successful
Job opportunities	Excellent, Driven, Mindset, International, Lacking, Unique
Flexibility	Strict, Dependent, Tremendous, Encourage, Minimal, Great
Personal growth	Exponential, Poor, Constant, Hierarchy, Potential, Constrain
Employees /Co-workers	Excellent, Cooperative, Competent, Stagnant, Unfriendly, Pretend
Employee experience	Diverse, Useful, Firsthand, Horrific, Odd, International
Employee communication	Strong, Transparent, Cryptic, Remote, Awful, Effective
Office culture	Cooperative, Balanced, Exciting, Suffered, Abysmal, Bias
Company business	Professional, Booming, Structured Challenging, Competitive, Steady
Career development	Rewarding, International, Guided Challenging, Difficult, Solid

Table 1: Top 24 Features along with their respective features Phase.

2. Dataset collection

During our research, we utilized Glassdoor.com, a well-known online resource for the employment industry, to construct the dataset. The vast majority of the site's verified evaluations were from current or former employees and happy customers of the evaluated companies. Overall, the reviews reflect the company's good and poor aspects, as evaluated by the many raters. We relied heavily on this platform to collect datasets, confident in the high quality that would emerge from the reviewers' decision to remain nameless. It was found that the evaluations made use of a vocabulary that was both very professional and somewhat slangy. The time required for data preparation activities like normalization, cleansing, and tokenization decreased accordingly.

Our ensemble network's efficiency was greatly improved by the ELM model's usage of the large number of matches between these "clean" words and the lexicon utilised to create the word embeddings. Reviewers on Glassdoor.com are tasked with being fair and balanced in their assessments of companies by considering factors such as the firm's advantages and negatives. This method is so precise that we were able to collect a comprehensive dataset that is fair and includes both positive and negative information about the company. Expressions like "I have nothing to say/complain about here" are either false positives or false negatives in the pros/cons section. This is because the point of the pros/cons section is to do just that—point out the benefits and drawbacks. Therefore we advise readers to continue with care. Given the prevalence of fraudulent remarks online and offline, we reasoned it was fair to include them in our data collection. An official API from Glassdoor was used to compile the dataset.[5] Our list of sixty major companies spans several economic sectors, from information technology. Over 20,000 testimonials from many different fields we have collected thus far. As was previously indicated, the assessments also offer a synopsis of the company's strengths and weaknesses. Thanks to our profound comprehension, we could dissect each review into positive and negative parts. This allowed us to implement automatic review tagging. Despite this, we eliminated the possibility of any false positives by doing a human check after that. Reviews of Accenture have been overwhelmingly positive, with many workers applauding the company's "great fundamental beliefs," "wonderful people," and "brilliant people-oriented training

programmes." Our staff may get insight into various facets of the company by collecting data on factors such as compensation, benefits, and other incentives.[6]

Company	Reviews	Company	Reviews
Accenture	1000	HP	150
Adobe	998	HSBC Holdings	1850
Aeropostale	874	IBM	150
Aflac	368	Intel Corporation	998
Autodesk	752	Intuit	972
Bank of China	212	Marriot International	980
Booz Allen Hamilton	976	Microsoft	1000
Broadcom	151	Mosanto	396
Brocade	990	Morningstar	733
Camden Property	203	National Instruments	712
Capital One	997	NetApp	800
CarMax	959	Nordstorm	839
Chesapeake Energy	725	OCBC	268
Cisco	980	Paychex	916
Citibank	1852	Qualcomm	919
Colgate-Palmolive	776	Quest Global	150
Creative Technology	150	Rackspace Hosting	732
Darden Restaurants	994	Samsung	150
DBS	288	SCB	942
Devon Energy	336	Singtel	480
DreamWorks Animation	978	StarBucks	828
EOG Resources	127	Starhub	150
FactSet	976	Stryker	904
FedEx Corporation	850	SVB Financial Software	277
Flextronics	150	J.M. Smucker Company	318
General Mills	1036	Ultimate Software	524
Goldman Sachs	970	Umpqua Bank	242
Google	756	Union Overseas Bank	265
Hasbro	458	World Foods Market	757
Herman Miller	540	Yes Bank	176

Table 2: The total number of reviews received by each firm [5]

Note: Priority was given to include Glassdoor.com evaluations with the highest 'helpful' tag in the dataset. These evaluations were chosen to represent each of the featured companies.

Table 2 summarises all reviews written about each organization. The target number of testimonials was set at between 500 and 1,000. There were more than a thousand reviews available for some firms on Glassdoor, while there were less than 500 reviews for others.

2.1.Preprocessing

As was previously indicated, it does not matter who is performing the grading; all evaluations must comply with the same rigorous methodology. Instead, we used a random number generator with a high degree of pseudo-randomness to rearrange the data and eliminate the chance of any patterns. We used regular expressions to strip the text of all the #hashtags, #urls, and #links preventing it from being processed any further. Nonetheless, we have included some of the many smiley faces and other emoticons utilized throughout the assessments. We created a very complex, context-aware algorithm.

3. Backgrounds of the study:

3.1.Features-sentiment analysis

The research aims to improve upon an existing technique called Features-sentiment analysis, which gives a map of companies into an n-dimensional plane. For each company, an aggregate sentiment strength is assigned to various areas of the organization based on employee feedback (including salary, work life, and location). The collection of sentiment scores is represented mathematically as the vector where si is the sentiment score for aspect I and n is the number of

factors (s1; s2;::::; sn). After constructing these vectors for each company in our dataset, we could derive the companies' aspect-emotion analysis.

3.2.Doc2vec for review level analysis

The ELM module is one of the two paths that make up our ensemble-based aspect-sentiment analysis. That's because the ELM module is a two-way piece of hardware. As a prerequisite to implementing the ELM module, we must transform the raw text into review-ready analysis. For optimal results, utilize Doc2vec. [7] While word2vec was successful in its own right, paragraph2vec (also known as doc2vec) is a more advanced variant of the same basic idea algorithm. Word2vec is a popular programme that creates embeddings for words. With the CBOW architecture, a neural network may be trained to obtain word embeddings relevant to their context. In a sense, Word2vec may be considered an embedding in and of itself. The 300-dimensional vectors were prepared using 100 billion Google News words. Doc2vec prioritizes the unverified learning of nonstop representations for more enormous stretches of text, such as phrases, paragraphs, or even full texts. Since our dataset includes in-depth reviews, we conclude that Doc2vec is an appropriate method. With the gensim8-provided Python implementation, our team could extract 300-dimensional embeddings for use in feeding the ELM model with sentiment analysis and classification data. [8]

3.3.Sentiment Dictionary/Lexicons

To incorporate polarity-based scoring for attributes, we created a lexicon of sentences that contained those polarities and used it in our ensemble method. The OED and SentiWordNet were our primary resources. [9] and entice them to develop this dictionary.[10] A strict threshold of 0.25 was maintained for the polarity scores of the words across both resources to filter out noisy phrases that don't contribute enough polarity to the aspects. This allowed us to remove unnecessary padding from the text. SentiWordNet was prioritized above SenticNet when the same word existed in both databases; the term polarity pair was selected from SentiWordNet. As far as possible, we avoided re-stating facts that were already well known. Following the dictionary's development, we included a lookup table containing its definitions in our algorithm (details of which are given below).

3.4.Features' Level Sentiment Analysis

Opinion mining experts have been presented with various analytical granularities, each with its advantages and disadvantages.[11] Mining of opinions depending on aspects. [12] It pays special attention to how a substance's polarity affects the characteristics it exhibits. Features are a lens through which one may examine the concepts represented in the available language. This concept is sometimes referred to as an opinion target. By saying something like, "The screen of my phone is quite nice, and its resolution is fantastic," the reviewer is expressing a favourable assessment of the gadget. However, the focus of the review is on display and pixel density. Features extraction is separating individual components of an opinion piece of writing apart. Two types of facets may be defined when talking about features-based opinion mining: explicit and implicit. Words or phrases that make the readership of a little piece very evident are examples of "explicit components." For example, "screen" and "resolution" are extensively explored as opinion targets in the sample above. An implicit aspect, on the other hand, is a word that indicates the opinion purpose of an opinionated document without explicitly saying so. This component was not mentioned explicitly in the previous sentence. To display the same thing, you might also say, "This camera is sleek and very affordable," which would be as accurate.

4. Extreme learning machine (ELM)

The data were categorized using ELM, a supervised classification approach. ELM-entitled methodology [13] when this solution was implemented, problems with the network's back-propagation stopped occurring. [14] training, Think about things like the possibility of slow convergence rates, the calibration of optimization parameters, and anything else that could be significant. In addition, local minima may form, calling for retraining or several attempts. A scalar

output indicates that the network has only one output unit without sacrificing generalizability, which is necessary for the ELM learning problem settings which require a training set, X, of N labelled pairings using the equation (xi; Yi), where xi 2 Rm is the i-th input vector and yi 2 R is the predicted 'target' value. This is because a training set, X, of N labelled pairings using the equation is required for the ELM learning problem settings. The m neurons in the input layer are linked to the Nh neurons in the hidden layer through weights written f WJ 2 Rm; j = 1;:::; Nhg. Because it incorporates a bias factor denoted by b j and a nonlinear 'activation' function characterized by the jth) hidden neuron's response to an input is nonlinear (_).

$$a_j(\mathrm{x}) = \varphi(\hat{\mathrm{w}}_j \cdot \mathrm{x} + \hat{b}_j) \qquad (1)$$

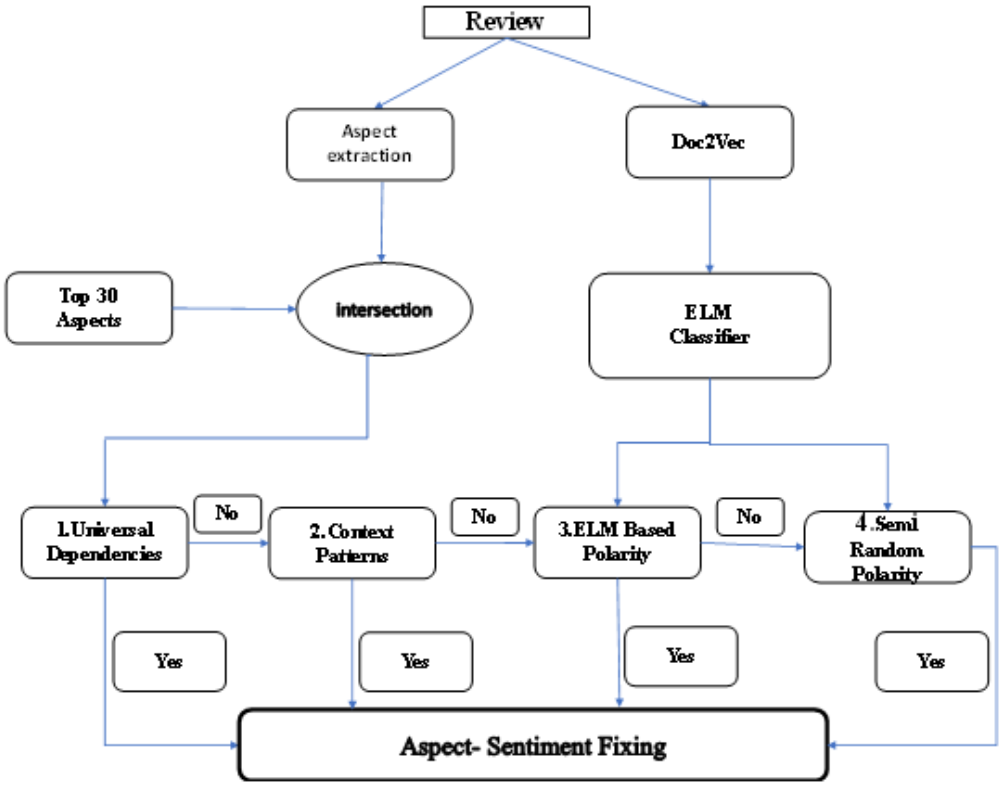

Fig. 2: The flowchart of the algorithm.

Remember that (1) may include a more extensive range of applications. [16]

This, however, is irrelevant to the analysis that follows. Using a weighted connection vector defined by _wj 2 RNA, a group of hidden neurons are linked to the output neuron without creating bias. The following formulas make up the network's aggregate output function, which is represented by f(x):

$$f(\mathrm{x}) = \sum_{j=1}^{N_h} \bar{\mathrm{w}}_j a_j(\mathrm{x}) \qquad (2)$$

It is common to practise generating an "activation matrix" H with the following values for the entry fhij 2 H: I = 1;:::; N; j = 1;:::; The jth hidden neuron's activation value, denoted by the symbol NHG, is determined by the ith input pattern. The elements of the H matrix are:

$$\mathrm{H} \equiv \begin{bmatrix} \varphi(\hat{\mathrm{w}}_1 \cdot \mathrm{x}_1 + \hat{b}_1) & \cdots & \varphi(\hat{\mathrm{w}}_{N_h} \cdot \mathrm{x}_1 + \hat{b}_{N_h}) \\ \vdots & \ddots & \vdots \\ \varphi(\hat{\mathrm{w}}_1 \cdot \mathrm{x}_N + \hat{b}_1) & \cdots & \varphi(\hat{\mathrm{w}}_{N_h} \cdot \mathrm{x}_N + \hat{b}_{N_h}) \end{bmatrix} \qquad (3)$$

The values of f WJ; b JG in (1) are fixed in stone in the ELM model. The only degrees of freedom are in the digits f, w, and j; b, and g are in (2). Reducing the convex cost optimally is the solution to the training issue.

$$\min_{\{\bar{\mathbf{w}},b\}} \left\|\mathbf{H}\bar{\mathbf{w}} - \mathbf{y}\right\|^2 \tag{4}$$

A matrix pseudo-inversion may be used to find the only L2 solution.

$$\bar{\mathbf{w}} = \mathbf{H}^{+}\mathbf{y} \tag{5}$$

A simple and successful method for training an ELM entails the following steps:

a. The input weights (wi) and biases (b I) of all the hidden neurons will be created arbitrarily.
b. Using equation (3) as a reference, we may determine the activation matrix H.
c. The output weights may be obtained by solving a pseudo-inverse problem, as explained in (5).

Despite the seeming ease of the ELM method, the most significant finding is that networks with random weights in the hidden layer have a remarkable ability to represent data. [17] According to the idea in [18], regularisation techniques may help improve its generalization performance as well. As can be seen in the following example, this results in an L2 regularisation factor being applied to the cost function (4)

$$\min_{\bar{\mathbf{w}}} \left\{\left\|\mathbf{H}\bar{\mathbf{w}} - \mathbf{y}\right\|^2 + \lambda\left\|\bar{\mathbf{w}}\right\|^2\right\} \tag{6}$$

4.1.Algorithmic specifics for ensemble architecture

In this study, we provide a hybrid method for tagging reviews with emotions. This approach combines unsupervised learning with machine learning. The dependent structure of the evaluation determines the initial polarity assigned to each item in the study. This strategy only works if the aspect term is linked to a phrase that is both polar and present in the emotional lexicon. If there are no opposite terms ociated with the aspect word in the emotion dictionary, we will use supervised classifiers, most notably ELM.

Unlike other machine learning models, such as support vector machines, ELM may help you reduce time and effort throughout the model-building process without sacrificing accuracy. With the likely ultimate deployment in an online and real-time situation, we spend a lot of effort on training. Figure 2 is a flowchart depiction of the proposed procedure.

4.2.Features Extraction

Our method for extracting aspects combined the Convolutional Neural Network and the language patterns described in [19]. This is the set of characteristics we were able to glean:

- Employment is an umbrella phrase that encompasses a wide range of job titles, responsibilities, and other aspects of working for a corporation—having "great people" and "stable and secure employment" as two of the best things about my job are only one example.

- Employees/Colleagues, in a nutshell, reveals the calibre of the company's staff and the quality of its relationships with its workers. This is epitomized by a successful company that has fostered a high-performance culture and attracted a pool of skilled workers.
- Work hours and patterns may be deduced from the positive company culture, as seen by the highly generous vacation and time off policies and the regular sabbaticals for full-time staff.

- Management, an in-depth look at the company's highest echelons; "Great place to work," for instance. A management style that takes into account the needs of all employees. Fantastic accomplishment.
- The customs and traditions of an organization are sometimes referred to as the "workplace culture" to simplify the discussion. Care is taken with the procedures, policies, culture, and staff—enormous benefits.
- "Competitive salary, Nice location, Full freedom" are just a few comments on the company's location that can be found under "Location."
- In this context, "job life" refers specifically to the type and quality of one's employment, as well as the work-life balance given by the organization, such as a friendly working atmosphere, convenient location, and exciting duties.
- Compensation, whereby pay-related remarks like "Wage is Fine" provide an in-depth look at the salary ranges offered by the organization. Great perk...until fights break out over who gets the most food. Much time is spent relaxing (which results in a lack of inventiveness).
- Salary, bonuses, and other benefits that are appropriate for this position are detailed below. These perks also change depending on one's level of seniority inside the organization. Vacation time, stock options, and a fitness centre membership with no cost to the employee is just some of the perks given.
- An excellent workplace has numerous components, including management that cares about workers' careers and a friendly atmosphere.
- The term "employee experience" refers to the perspectives of employees of varying skill levels and the range of experiences that might be received via working for the company, including but not limited to: It may be difficult for new hires, regardless of their degree of experience, to have their opinions taken seriously in a large organization.
- There is an official staff area where problems with hiring, firing, and retention may be addressed.
- Training received while working is known as "on-the-job training," often provided by an employer. For instance, you may need to adjust your schedule or be unavailable for a specific time. subpar schooling
- Because it is a complex company and constantly evolving, success takes time to realize. Personal and professional growth possibilities. Organizational growth.
- Effectiveness in Leadership, which examines the level of motivation and efficiency shown by top-level officials and the organization's leadership.
- Leadership is a study of influential leaders and their functions. Management has no idea how to make the most of its seasoned staff.
- Others who want power inspire those who desire it via their relationships with one another in politics.
- For instance, an inquiry into the firm's business may be inferred from the statement "Business is thriving at the company," which alludes to the success and trade of the company.
- Predictions about one's professional development and the future of one's organization, such as plans to become global in two years.
- In this context, "vacation" refers to the number of paid days off that the company provides the workers, as in "Paid vacation during the summer." The ability to freely traverse international boundaries.
- You may tell how supportive an organization is by observing its members' interactions with one another, such as how supervisors care for their direct reports.
- A company's dedication to a flexible work environment may be gauged by how much remote work is an option for workers.
- The performance of an employee is defined as their activities in the workplace, whether it be the quality and quantity of their labour or the demonstration of exceptional talents.
- Respect on the job is how well employees think their managers and colleagues are doing their jobs.

- Projects at work, corporate projects, products, and corporate strategies are covered; for instance, the company provides a broad range of products and a substantial number of projects.
- The strength of the market in which a business works is reflected in the firm's ability to succeed there.
- This section examines the firm's technical underpinnings, including that most of the firm's equipment is built on obsolete systems.
- Most machines, for instance, are malfunctioning, drawing attention to the fact that the company must overcome operational, regulatory, and internal obstacles.
- Both required and elective skills for the organization are outlined by the knowledge scope. Examples include "This process requires expert expertise," etc.
- Employee communication is the two-way flow of information between employees and their respective employers. Saying something like, "People at this company are extremely engaged," is an example of effective internal communication.
- Stress in the workplace, economic factors (such as low salaries), and the upsides (increased productivity) of pressure in a demanding business environment.

Features	Persistence
Employees/Co-workers	7659
Work Life	7305
Perks/Benefits	4565
Office culture	4192
Working time	3658
Salary	3323
Management	2654
Job Opportunities	2129
Employee experience	1288
Location	627
Leadership	599
Technology	490
Politics	451
Flexibility	340
Company business	116

Table 3: Extracted features with their corpus Persistence.

4.3.Contradiction Assignment for the Features

In this section, we'll examine how the polarity of each variable is established. Conditional Modifiers That Can Be Used Anywhere Our final objective is to find the universal dependencies of the included features to calculate the polarity score for each of the 30 most significant elements of a review. Our main study focuses on the interplay between adjective modifiers, adverbial modifiers, and nominal subjects (subj). To help, we provide snippets of the reviews included in our dataset. These images attempt to demonstrate the interdependence of these parts.

- Many opportunities to move forward in one's field. Plenty of possibilities. They have conservative political views. Extremely out of date and rooted in the past. Adv med (political, Very)
- Meeting interesting new individuals is an excellent advantage of work trips. Subj (travelling, perks)

The StanfordCoreNLP Parser is the software we use to discover these shared features. Once a causal relationship has been established, its valence may be determined by examining the "trigger" terms. This aspect's mood is determined by the polarity of the modifiers. Therefore we look them up in our made-up sentiment dictionary. The contradictions upon which the aspect score is based provide a reliable assessment of the property in question. This is done for the 30 highest weighted criteria in the evaluation.

Dialects and their Variations The ensemble advances to the next stage when the trigger word is not included in their emotional vocabulary. Here, we'll take a step back and examine the larger picture (window size - 5 terms used, including the aspect). We utilize the dependent patterns mentioned in [20] as a basis for determining an overall polarity score for the aspect. In our view, highly polar

terms in the context will increase the polarity of the aspect as a whole. The disadvantages are now advantages, and they can be dealt with appropriately. A polarity value may be assigned to anything using ELM data. If none of these alternative techniques provides a suitable grade for the component, we shall resort to the result reached by our ELM model (1 - positive, 0 - negative). Using the results from the ELM, we change the aspect word by first looking it up in the sentiment dictionary and then using the following algorithm.

$$aspect_{score} = (e_{out}) * (lookup(aspect)) + (1 - e_{out}) * (-1 * lookup(aspect))$$

Here, e out is the expected output of the ELM for the review, and lookup(aspect) is the polarity score searched up in the sentiment dictionary for the aspect term. Randomized ELM score when an aspect word is not included in the sentiment dictionary, a polarity value is arbitrarily generated using the ELM's output. The formula for the random score is as follows:

$$aspect_{score} = \begin{cases} rand(0,1) & , e_{out} \equiv 1 \\ rand(-1,0) & , e_{out} \equiv 0 \end{cases}$$

To generate random real numbers between [a,b], the rand(a,b) function uses a random number generator. It was only in 2% of the occasions where we had to employ a random number generator to assign weights to the various characteristics. These findings demonstrate that the semi-random production of aspect polarities did not affect the overall aspect-sentiment embeddings. Still, we think being fully automated is the way to go for our future endeavours. Therefore we plan to do just that.

5. Features-Sentiment Analysis

In the last part, we examined how to discern each characteristic's direction. Then, we averaged the ratings of the polarity related to that element throughout all assessments of the organization to get an overall grade for each factor. If we treat each facet as a single dimension, and if the projection of a firm along a given size is equal to the polarity value of a firm in that facet, then we may project each business.

In Table 4, Cosine similarity scores compare businesses in the same and distinct fields. Even though both Goldman Sachs and DBS provide financial services, their similarity score was lower than that of other banks. DBS and SCB (Standard Chartered Bank) have ratings that are substantially closer to one another.

Table 5 Ranking the best and not best places to work in the information technology and financial sectors based on employee feedback on perks, compensation, and corporate culture. Pay satisfaction is high among IT and finance employees, as seen in Figures 3b and 3a. The great majority of bank employees, however, had unfavourable impressions of their company's culture.

Company A	Company B	Cosine Similarity
Accenture	Booz Allen Hamilton	0.548
Accenture	FedEx	0.733
Google	Microsoft	0.549
Microsoft	Intel	0.486
Adobe	HP	0.370
Adobe	Google	0.276
Adobe	IBM	-0.214
OCBC	Goldman Sachs	0.422
Goldman Sachs	DBS	-0.098
DBS	SCB	0.313
Goldman Sachs	SCB	0.041
Singtel	Broadcom	0.424
Singtel	Starhub	-0.244
Microsoft	Stryker	0.687
National Instruments	Microsoft	-0.117
NetApp	Hasbro	0.655
Monsanto	Quest GLobal	-0.380

Table 4: Calculated Cosine Similarities between the Companies Using Their Feature-Sentiment Analysis

	Location		Salary		Work Life	
	Tech	Finance	Tech	Finance	Tech	Finance
Best	Microsoft Intel Adobe Google HP	Umpqua Bank Goldman S SCB Citibank HSBC	Intel Adobe Microsoft Cisco Google	Yes Bank Goldman S HSBC OCBC Citibank	Adobe Microsoft Google Cisco FactSet	Goldman S Bank of China SCB UOB HSBC
Not Best	IBM Creative NetApp Cisco FactSet	Yes Bank OCBC DBS UOB Bank of China	Creative Flextronics FactSet Samsung HP	UOB Bank of China SCB Umpqua Bank DBS	Samsung HP NetApp Creative Intuit	Citibank DBS OCBC Umpqua Bank Yes Bank

Table 5: In the business worlds of technology and finance, we look at which companies provide the most significant and worst compensation and workplace cultures.

Conclusion:

In this research, we examined the process of dynamic embeddings in the context of business. We conducted a sentiment analysis on Glassdoor reviews of several companies, focusing on hitherto unexplored aspects. These findings are explained in a variety of practical ways. We surveyed staff on various topics, including how they felt about their compensation, perks, and general satisfaction with their work environment. The results of this study might be used by businesses to boost morale by helping management identify and address issues that matter to their staff. Conversely, this information will aid job-seekers in identifying leading organizations in their chosen area. Future

work on a user-product sentiment model will concentrate mainly on using the current user rating as input. Further exploration of the feeling from several angles will be necessary for the next step.

References:

1. Rajiv Ratn Shah. Multimodal analysis of user-generated content in support of social media applications. In Proceedings of the International Conference on Multimedia Retrieval (ICMR), pages 423–426. ACM, 2016.
2. Rajiv Ratn Shah, Yi Yu, Akshay Verma, Suhua Tang, Anwar Dilawar Shaikh, and Roger Zimmermann. Leveraging multimodal information for event summarization and concept-level sentiment analysis. Proceedings of the Knowledge-Based Systems (KBS), pages 1–8, 2016.
3. Rajiv Ratn Shah, Yi Yu, and Roger Zimmermann. Advisor: Personalized video soundtrack recommendation by late fusion with heuristic rankings. In Proceedings of the International Conference on Multimedia (MM), pages 607–616. ACM, 2014.
4. Andy Moniz and Franciska de Jong. Sentiment analysis and the impact of employee satisfaction on firm earnings. In European Conference on Information Retrieval, pages 519–527. Springer, 2014.
5. https://www.glassdoor.com/developer/index.htm
6. Andrea Esuli and Fabrizio Sebastiani. Sentiwordnet: A publicly available lexical resource for opinion mining. In Proceedings of LREC, volume 6, pages 417–422, 2006.
7. Quoc V Le and Tomas Mikolov. Distributed representations of sentences and documents. In ICML, volume 14, pages 1188–1196, 2014.
8. Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781, 2013.
9. Andrea Esuli and Fabrizio Sebastiani. Sentiwordnet: A publicly available lexical resource for opinion mining. In Proceedings of LREC, volume 6, pages 417–422, 2006.
10. Erik Cambria, Soujanya Poria, Rajiv Bajpai, and Björn Schuller. SenticNet 4: A semantic resource for sentiment analysis based on conceptual primitives. In COLING, 2016.
11. E. Cambria, B. Schuller, B. Liu, H. Wang, and C. Havasi. Statistical approaches to concept-level sentiment analysis. IEEE Intelligent Systems, 28(3):6–9, May 2013.
12. Minqing Hu and Bing Liu. Mining and summarizing customer reviews. In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, pages 168–177. ACM, 2004.
13. Guang-Bin Huang, Dian Hui Wang, and Yuan Lan. Extreme learning machines: a survey. International Journal of Machine Learning and Cybernetics, 2(2):107–122, 2011.
14. Sandro Ridella, Stefano Rovetta, and Rodolfo Zunino. Circular backpropagation networks for classification. Neural Networks, IEEE Transactions on, 8(1):84–97, 1997.
15. Thomas P Vogl, JK Mangis, AK Rigler, WT Zink, and DL Alkon. Accelerating the convergence of the back-propagation method. Biological cybernetics, 59(4-5):257– 263, 1988.
16. Guang-Bin Huang, Lei Chen, and Chee-Kheong Siew. Universal approximation using incremental constructive feedforward networks with random hidden nodes. Neural Networks, IEEE Transactions on, 17(4):879–892, 2006.
17. Guang-Bin Huang, Dian Hui Wang, and Yuan Lan. Extreme learning machines: a survey. International Journal of Machine Learning and Cybernetics, 2(2):107–122, 2011.
18. Guang-Bin Huang, Hongming Zhou, Xiaojian Ding, and Rui Zhang. Extreme learning machine for regression and multiclass classification. Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, 42(2):513–529, 2012.
19. Soujanya Poria, Erik Cambria, and Alexander Gelbukh. Aspect extraction for opinion mining with a deep convolutional neural network. Knowledge-Based Systems, 2016.
20. Soujanya Poria, Basant Agarwal, Alexander Gelbukh, Amir Hussain, and Newton Howard. Dependency-based semantic parsing for concept-level text analysis. In Computational Linguistics and Intelligent Text Processing, pages 113–127. Springer, 2014.

Books:

1. Agarwal, Basant, and Namita Mittal. *Prominent feature extraction for sentiment analysis*. Berlin: Springer International Publishing, 2016.
2. Agarwal, Basant, et al., eds. *Deep learning-based approaches for sentiment analysis*. Singapore: Springer, 2020.
3. Poria, Soujanya, Amir Hussain, and Erik Cambria. *Multimodal sentiment analysis*. Cham: Springer International Publishing, 2018.
4. Pozzi, Federico, et al. *Sentiment analysis in social networks*. Morgan Kaufmann, 2016.
5. Rajput, Dharmendra Singh, Ramjeevan Singh Thakur, and S. Muzamil Basha, eds. *Sentiment Analysis and Knowledge Discovery in Contemporary Business*. IGI Global, 2018.
6. Keenan, Mark JS. *Advanced positioning, flow, and sentiment analysis in commodity markets: bridging fundamental and technical analysis*. John Wiley & Sons, 2020.
7. Liu, Bing. *Web data mining: exploring hyperlinks, contents, and usage data*. Vol. 1. Berlin: springer, 2011.
8. Chaudhuri, Arindam. *Visual and text sentiment analysis through hierarchical deep learning networks*. New York: Springer, 2019.
9. Deng, Li, and Yang Liu, eds. *Deep learning in natural language processing*. Springer, 2018.
10. Satapathy, Ranjan, Erik Cambria, and Amir Hussain. *Sentiment Analysis in the Bio-Medical Domain*. 2017.