IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A Study On Compensation Analytics Using Machine Learning Techniques In Cholamandalam Investment And Finance Company Limited

[1]Sneha G, [2]Ms.J. Sofia vincent

[1]Student ,[2]Professor

[1]Department Of Management Studies

[1]Panimalar Engineering College, Chennai, India

***Abstract*—** This study explores the application of machine learning techniques in compensation analytics at Cholamandalam Investment and Finance Company Limited (CIFCL). As organizations increasingly shift toward data-driven decision-making, this research investigates how predictive models can enhance fairness, transparency, and efficiency in compensation management. Using a dataset of employee compensation records, the study applies regression algorithms, feature selection methods, and dimensionality reduction techniques such as PCA to identify key factors influencing salary structures. Polynomial Regression emerged as the most accurate model for predicting compensation, while analysis also revealed strong correlations among salary components and the potential for multicollinearity. The findings underscore the value of machine learning in optimizing pay structures, detecting anomalies, and supporting strategic HR decisions. Despite some limitations related to data size and model complexity, the study demonstrates that AI-driven compensation analytics can significantly improve employee retention, equity, and organizational performance.

***KEY WORDS*—** Compensation Analytics, Principal Component Analysis (PCA), Multicollinearity

I. INTRODUCTION :

In the contemporary business environment, compensation management has evolved beyond traditional payroll systems to become a strategic tool for talent acquisition, employee retention, and performance optimization. As companies strive to build competitive and high-performing workforces, the need for data-driven compensation strategies has become increasingly evident. The intersection of compensation analytics and machine learning presents an innovative frontier, enabling organizations to gain actionable insights into pay structures, performance linkages, and workforce equity. This project aims to explore how machine learning techniques can enhance compensation analytics within Cholamandalam Investment and Finance Company Limited (CIFCL), a leading financial services provider in India.

II. NEED OF THE STUDY :

- Machine learning enables efficient analysis of large volumes of compensation data, helping businesses identify trends and anomalies for informed pay decisions.
- By comparing internal salary trends with industry standards, ML helps optimize compensation structures, ensuring fairness and boosting employee morale.
- ML can predict potential employee turnover by analyzing pay patterns and satisfaction levels, enabling proactive retention strategies through timely interventions.

OBJECTIVES OF THE STUDY :

PRIMARY OBJECTIVE

- To evaluate the Compensation analytics using machine learning techniques Cholamandalam Investment and Finance Company Limited.

SECONDARY OBJECTIVES

- To analyze the effectiveness of machine learning techniques in compensation analytics
- To evaluate the accuracy and predictive power of different machine learning algorithms in forecasting employee compensation trends.
- To assess the impact of AI-driven compensation models on employee satisfaction, retention, and productivity.
- To compare traditional compensation analysis methods with machine learning-based approaches in terms of efficiency and fairness.
- To provide strategic recommendations for implementing machine learning in compensation management.

III. SCOPE OF THE STUDY :

- This study focuses on the application of machine learning techniques in analyzing and optimizing compensation structures at Cholamandalam Financial Holding Ltd. It aims to evaluate how data-driven compensation strategies can improve workforce management within the NBFC sector. The findings may also be applicable to similar financial institutions looking to enhance their salary and benefits frameworks.
- The study explores various machine learning algorithms such as regression models, decision trees, and neural networks to analyze compensation data. It examines how AI can detect salary trends, predict employee retention risks, and ensure fair pay distribution. The research will also assess the role of data analytics tools in automating salary decisions and forecasting future compensation needs.

IV. LIMITATIONS OF THE STUDY :

V. The accuracy of machine learning models depends on the availability and quality of compensation-related data. Incomplete, inconsistent, or biased data can affect the reliability of insights and predictions, leading to potential inaccuracies in salary analysis and forecasting.

VI. The study primarily focuses on salary structures, incentives, and financial benefits**,** but non-monetary factors like work culture, job satisfaction, career growth, and employee well-being are not deeply analyzed. These factors also play a crucial role in employee retention and satisfaction.

VII. REVIEW OF LITERATURE :

VIII. **Thompson (2023)** reported that machine learning models using historical salary data could accurately predict future compensation trends. The study highlighted that using AI-powered forecasting tools helped organizations stay competitive by ensuring that salary increments aligned with market conditions.

IX. **Singh and Kapoor (2023)** investigated how AI-powered compensation analytics reduced wage disparities and enhanced employee motivation. Their study found that ML-driven compensation systems helped HR teams ensure pay equity across different roles and departments.

X. **Feng, Liu, & Yin (2023)** compared deep learning and conventional machine learning algorithms for salary prediction. Utilizing a dataset from Kaggle, they developed models using Convolutional Neural Networks (CNN) and Random Forest algorithms. The study found that the CNN model outperformed the Random Forest in terms of accuracy and stability, suggesting that deep learning approaches can yield better results in salary prediction tasks.

XI. RESEARCH METHODOLOGY :

Research methodology is that rational approach used by researchers to investigate a particular topic or problem. It basically involves certain steps and techniques carried out in collecting,analysing, and interpreting data to answer research questions or achieve desired research objectives. It basically outlines the general design of the study pertaining to the type of research to be conducted, the methods of data collection, the sampling strategy, and the technique of data analysis.

QUANTITATIVE RESEARCH

The type of research used in the study is Quantitative research .Quantitative research is a method that focuses on quantitative measurement of data collection and analysis.

SOURCE OF DATA COLLECTION

Secondary data refers to information that has already been collected, processed, and made available by someone else for a purpose other than the current research. It can come from sources like government reports, company records, books, articles, online databases, and past research studies.

TOOLS USED

- IBM SPSS
- EXCEL

IBM SPSS(Statistical Package for the Social Sciences): A widely used program for statistical analysis in social science. It offers advanced statistical procedures and descriptive analysis for data interpretation.

EXCEL: Microsoft Excel was used for data organization, analysis, and visualization. It facilitated data entry, computation of descriptive statistics, and the creation of charts and graphs to support the findings of the study.

STATISTICAL TOOLS AND TECHNIQUES

1. Descriptive Statistics
2. Correlation Matrix Analysis

I. DATA ANALYSIS AND INTERPRETATION :

1.Descriptive Analysis

Descriptive statistics								
		Basic Pay	Health Insurance	Retirement Plans	Travel Allowance	Childcare Assistance	Bonus	Incentive
N	Valid	2654	2654	2654	2654	2654	2654	2654
	Missing	0	0	0	0	0	0	0
Mean		946194.4	44235.08	12459.75	10680.01	17799.72	5.34E4	5529.28
Median		900000.0	50199.00	13641.00	11692.00	19487.00	5.85E4	6275.00
Mode		700000	65258	15005	12862	21436	64308	8157
Std. Deviation		4.648E5	23049.783	2958.922	2536.465	4227.191	1.268E	2.881E3
Variance		2.161E1	5.313E8	8755221.85	6433653.1	1.787E7	1.608E	8.301E6
Skewness		.003	-.393	-.622	-.622	-.622	-.622	-.393
Std. Error of Skewness		.048	.048	.048	.048	.048	.048	.048
Kurtosis		-1.236	-1.585	-1.216	-1.216	-1.216	-1.216	-1.585
Std. Error of Kurtosis		.095	.095	.095	.095	.095	.095	.095
Range		1500000	57258	8005	6862	11436	34308	7157
Minimum		200000	8000	7000	6000	10000	30000	1000
Maximum		1700000	65258	15005	12862	21436	64308	8157
Sum		3.E9	11739989	33068169	28344737	47240445	1.E8	1.E7

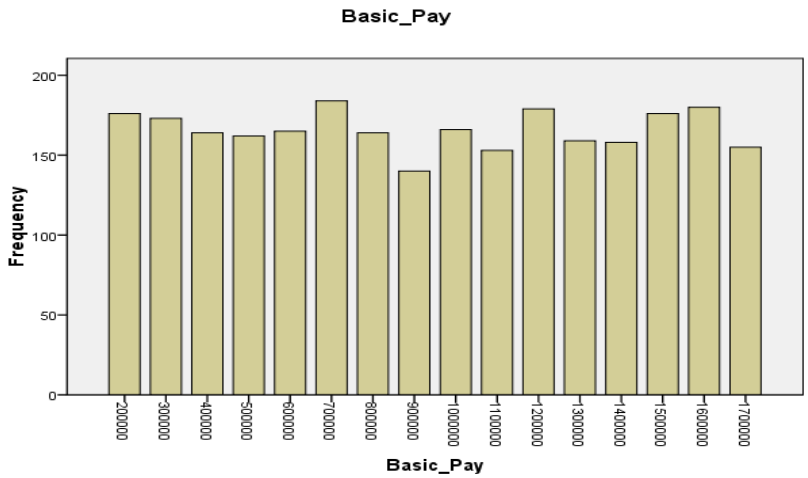
INTERPRETATION

The descriptive statistics show that all compensation and benefit variables Basic Pay, Health Insurance, Retirement Plans, Travel Allowance, Childcare Assistance, Bonus, and Incentive were reported by all 2,654 respondents with no missing data. The mean Basic Pay was ₹946,194.42 with a standard deviation of ₹464,800, while the median was ₹900,000 and the mode was ₹700,000. The highest frequencies in all benefit categories were concentrated at their maximum values (e.g., 50% received ₹65,258 in Health Insurance, ₹15,005 in Retirement Plans, and ₹8,157 in Incentives). All variables displayed negative skewness and platykurtic distributions, indicating that most values were concentrated towards the higher end with flatter tails. Standard deviations and variances showed varying levels of dispersion, with Basic Pay and Bonus exhibiting relatively higher variability. The range was also significant across components, especially for Basic Pay (₹1,500,000).

Table showing the basic pay

Basic_Pay					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	200000	176	6.6	6.6	6.6
	300000	173	6.5	6.5	13.1
	400000	164	6.2	6.2	19.3
	500000	162	6.1	6.1	25.4
	600000	165	6.2	6.2	31.7
	700000	184	6.9	6.9	38.6
	800000	164	6.2	6.2	44.8
	900000	140	5.3	5.3	50.0
	1000000	166	6.3	6.3	56.3
	1100000	153	5.8	5.8	62.1
	1200000	179	6.7	6.7	68.8
	1300000	159	6.0	6.0	74.8
	1400000	158	6.0	6.0	80.7
	1500000	176	6.6	6.6	87.4
	1600000	180	6.8	6.8	94.2
	1700000	155	5.8	5.8	100.0
	Total	2654	100.0	100.0	

Chart showing the basic pay



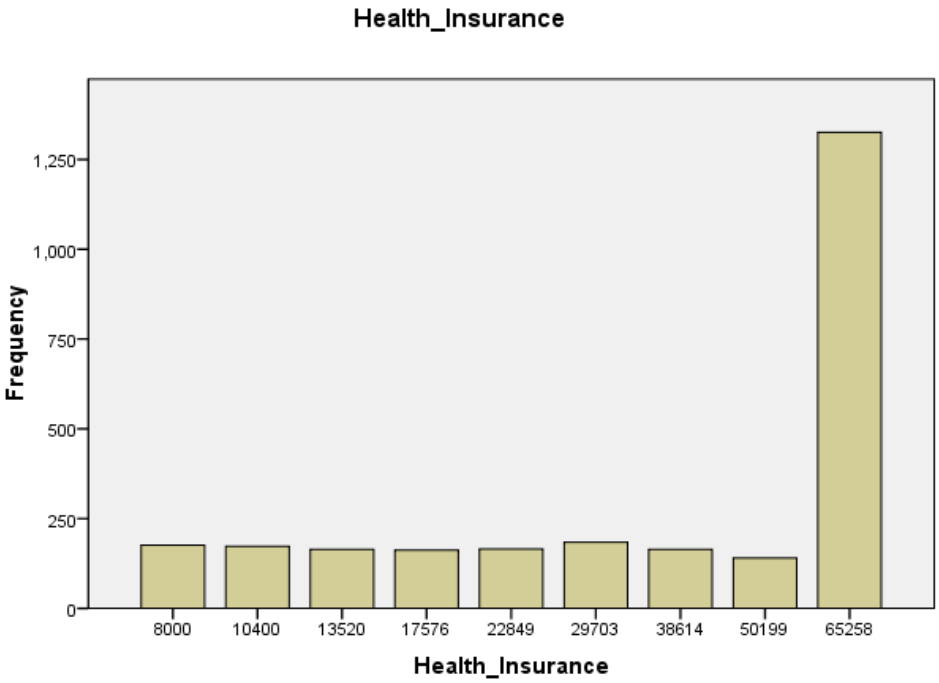
INTERPRETATION

The distribution of basic pay among employees at CIFCL reveals a relatively balanced salary structure, with most salary bands (ranging from ₹200,000 to ₹1,700,000) contributing between 5.3% and 6.9% of the total employee count. The most common basic pay levels are ₹700,000 (6.9%), ₹1600,000 (6.8%), and ₹1200,000 (6.7%), indicating slight clustering around mid-to-high salary brackets. The least common is ₹900,000 (5.3%), though the variation across all bands is minimal. This even spread suggests a standardized approach to compensation with no extreme skews, reflecting consistency in the organization's pay structure.

Table showing the health insurance

Health Insurance					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	8000	176	6.6	6.6	6.6
	10400	173	6.5	6.5	13.1
	13520	164	6.2	6.2	19.3
	17576	162	6.1	6.1	25.4
	22849	165	6.2	6.2	31.7
	29703	184	6.9	6.9	38.6
	38614	164	6.2	6.2	44.8
	50199	140	5.3	5.3	50.0
	65258	1326	50.0	50.0	100.0
	Total	2654	100.0	100.0	

3.2.1.2 Chart showing the health insurance



INTERPRETATION

The distribution of health insurance amounts among employees at CIFCL shows a significant concentration in one particular bracket. While the lower insurance values (ranging from ₹8,000 to ₹50,199) are relatively evenly distributed each accounting for around 5% to 7% of the total ₹65,258 stands out, covering 50% of the employees (1,326 out of 2,654). This suggests that half the workforce is enrolled in a standardized or upgraded health insurance plan, possibly due to policy changes or eligibility criteria. The uniformity across lower tiers and the sharp increase at the highest tier indicate a structured shift in benefits, reflecting organizational efforts to enhance employee welfare or streamline insurance offerings.

2.Correlation

Null Hypothesis (H_0): There is no statistically significant relationship between Basic Pay and other compensation components, namely Incentive, Bonus, Childcare Assistance, Travel Allowance, Retirement Plans, and Health Insurance.

Alternative Hypothesis (H_1): There is a statistically significant relationship between Basic Pay and other compensation components, namely Incentive, Bonus, Childcare Assistance, Travel Allowance, Retirement Plans, and Health Insurance.

Correlations

		Basic Pay	Health Insurance	Retirement Plans	Travel Allowance	Childcare Assistance	Bonus	Incentive
Basic Pay	Pearson Correlation	1	.932**	.925**	.925**	.925**	.925**	.932**
	Sig. (2-tailed)		.000	.000	.000	.000	.000	.000
	N	2654	2654	2654	2654	2654	2654	2654
Health Insurance	Pearson Correlation	.932**	1	.991**	.991**	.991**	.991**	1.000**
	Sig. (2-tailed)	.000		.000	.000	.000	.000	.000
	N	2654	2654	2654	2654	2654	2654	2654
Retirement Plans	Pearson Correlation	.925**	.991**	1	1.000**	1.000**	1.000**	.991**
	Sig. (2-tailed)	.000	.000		.000	.000	.000	.000
	N	2654	2654	2654	2654	2654	2654	2654
Travel Allowance	Pearson Correlation	.925**	.991**	1.000**	1	1.000**	1.000**	.991**
	Sig. (2-tailed)	.000	.000	.000		.000	.000	.000
	N	2654	2654	2654	2654	2654	2654	2654
Childcare Assistance	Pearson Correlation	.925**	.991**	1.000**	1.000**	1	1.000**	.991**
	Sig. (2-tailed)	.000	.000	.000	.000		.000	.000
	N	2654	2654	2654	2654	2654	2654	2654
Bonus	Pearson Correlation	.925**	.991**	1.000**	1.000**	1.000**	1	.991**
	Sig. (2-tailed)	.000	.000	.000	.000	.000		.000
	N	2654	2654	2654	2654	2654	2654	2654
Incentive	Pearson Correlation	.932**	1.000**	.991**	.991**	.991**	.991**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	
	N	2654	2654	2654	2654	2654	2654	2654

**. Correlation is significant at the 0.01 level (2-tailed).						

INTERPRETATION

The correlation analysis reveals a strong and statistically significant relationship between Basic Pay and other compensation components, including Health Insurance, Retirement Plans, Travel Allowance, Childcare Assistance, Bonus, and Incentive. Among these, Incentive and Health Insurance exhibit the highest correlation with Basic Pay, both showing a Pearson correlation coefficient of 0.932, indicating a very strong positive relationship. The other components Retirement Plans, Travel Allowance, Childcare Assistance, and Bonus each have a coefficient of 0.925, also signifying strong positive associations. All correlations are significant at the 0.01 level (2-tailed), suggesting that increases in these compensation components are associated with increases in Basic Pay. Given these statistically significant results, the null hypothesis which stated there is no relationship between Basic Pay and the other compensation elements is rejected. Thus, it can be concluded that there is a significant and positive relationship between Basic Pay and the various components of employee compensation.

FINDINGS

- The salary distribution at CIFCL is balanced, with most employees earning within the mid-to-high pay brackets.
- Half of the employees receive the highest health insurance benefit amount (₹65,258), indicating possible benefit standardization or policy changes.
- There is a strong positive correlation between Basic Pay and other compensation components such as Incentive, Bonus, Retirement Plans, Travel Allowance, and Childcare Assistance.
- Polynomial Regression proved to be the most accurate model for predicting employee compensation, showing the usefulness of non-linear modeling.
- Descriptive statistics reveal negative skewness and platykurtic distributions, indicating most employees receive benefits at the higher end of the range.
- Compensation components may show redundancy due to high correlation, requiring better structuring to ensure efficiency.
- The use of AI-driven models like Polynomial Regression is recommended to improve fairness, accuracy, and strategic planning in compensation.
- Techniques like Principal Component Analysis (PCA) should be applied to mitigate multicollinearity issues in compensation data.
- Future studies should consider non-monetary factors such as job satisfaction and work culture to provide a holistic understanding of compensation.
- Overall, the application of machine learning in compensation analytics can significantly enhance decision-making, employee equity, and organizational performance.

SUGGESTIONS :

- Implement standardized compensation structures to reduce variation and promote fairness across different employee levels.
- Integrate machine learning models, especially Polynomial Regression, into HR systems for accurate compensation forecasting and planning.
- Regularly evaluate compensation components to identify and eliminate redundancy caused by high inter-correlation among variables.
- Apply dimensionality reduction techniques such as Principal Component Analysis (PCA) to address multicollinearity and enhance model interpretability.

- Introduce continuous training for HR personnel on the use of AI and analytics tools to improve decision-making in compensation management.
- Expand data collection efforts to include qualitative factors like job satisfaction, career growth opportunities, and employee engagement.
- Benchmark internal compensation trends against industry standards to maintain competitiveness and retain top talent.
- Periodically review and revise health insurance and other benefits to ensure they align with employee needs and organizational goals.
- Promote transparency in compensation decisions through data-backed justifications, enhancing employee trust and morale.
- Encourage cross-functional collaboration between HR, data analytics, and finance teams to develop a holistic and strategic compensation framework.

CONCLUSION

The study effectively demonstrates the potential of machine learning techniques in transforming compensation analytics at Cholamandalam Investment and Finance Company Limited (CIFCL). By analyzing compensation data using tools such as Polynomial Regression and Principal Component Analysis, the research identified key trends, strong interrelationships between compensation components, and areas where standardization can enhance efficiency. The findings confirm that machine learning models can predict compensation accurately, support strategic HR planning, and promote fairness and transparency in pay structures. While data limitations exist, the study establishes that AI-driven approaches can significantly improve employee retention, organizational equity, and overall performance. Moving forward, integrating qualitative aspects such as job satisfaction and professional development will further enrich compensation strategies and ensure a more comprehensive understanding of employee needs.

BIBLIOGRAPHY

1. Thompson, M. (2023). Predicting salary trends with historical ML models. *Journal of Labor Economics and Technology*, **15**(4), 245–261.
2. Singh, R., & Kapoor, M. (2023). Using AI to reduce wage disparities. *International Journal of Human Capital and Information Technology Professionals*, **14**(3), 23–41.
3. Feng, Y., Liu, H., & Yin, W. (2023). Deep learning vs. machine learning for salary prediction. *Journal of Big Data*, **10**(1), Article 87.

Websites

- **https://matplotlib.org/**
- **https://scholar.google.com**
- **https://seaborn.pydata.org**