|  |
| --- |
| **Ain Shams University**  **Faculty of Computer & Information Sciences**  **Information System Department** |

**AutoInsight : Empowering companies with instant data analytics**

**This documentation submitted as required for the degree of bachelors in Computer and Information Sciences**

**By**

Maya Mohamed Mohamed [IS]

Farah Moataz Mohamed [IS]

Mazen Mostafa Hanafy [IS]

Mazen Raafat Abdel Hamed [IS]

Mayer Soliman Hedya [IS]

Bishoy Sedra Saber [IS]

**Under Supervision of**

**DR.Eman Amin**

[Supervisor Title],

Information System Department,

Faculty of Computer and Information Sciences,

Ain Shams University.

**TA.Verina Nashaat**

[Supervisor Title],

………………. Department,

Faculty of Computer and Information Sciences,

Ain Shams University.

**June 2025**

**Acknowledgements**

All praise and thanks to ALLAH, who provided me with the ability to complete this work. I hope to accept this work from me.

I am grateful of *my parents* and *my family* who are always providing help and support throughout the whole years of study. I hope I can give that back to them.

I also offer my sincerest gratitude to my supervisors, *Prof. Dr. ………………*, *Dr.*

*……………… and T.A Verina Nashaat* who have supported me throughout my thesiswith their patience, knowledge and experience.

**Abstract**

## In today's competitive business landscape, companies rely on **data-driven insights** to optimize performance and decision-making.

## However, many businesses face **layoffs, restructuring, or budget constraints**, limiting their ability to maintain dedicated data analytics teams.

## This creates a demand for **automated analytics solutions** that can bridge this gap and ensure continuous, reliable insights.

## To address this challenge, we developed **AutoInsight**, an intelligent **automated business analytics tool** that streamlines data processing and analysis.

## It automates **data preprocessing** and conducts **three types of analysis**: **diagnostic analysis** (identifying past trends and anomalies), **exploratory data analysis (EDA)** (uncovering patterns and relationships), and **predictive analysis** (forecasting future trends).

## Using **Facebook Prophet**, AutoInsight ensures accurate **trend detection, seasonality analysis, and anomaly detection** without manual intervention.

## Beyond forecasting, AutoInsight enhances user experience by offering **customizable reports** tailored to business needs. Additionally, it integrates a **chatbot powered by the Gemini API**, allowing users to interact through text and images for deeper insights into their dashboards.

## This feature helps users **interpret data, understand key metrics, and make informed decisions** efficiently.

## Results from real-world datasets demonstrate **high forecast accuracy** and **effective data-driven decision-making**. By automating analytics, businesses can **optimize performance, monitor financial trends, and enhance marketing, inventory, and pricing strategies**—all while minimizing manual effort.

## In conclusion, **AutoInsight provides a scalable, AI-driven solution** that enables companies to maintain **critical analytics capabilities** regardless of workforce constraints. It empowers businesses to make smarter decisions, enhance efficiency, and stay **resilient in a rapidly evolving market**.

## **الملخص**

## في بيئة الأعمال التنافسية اليوم، تعتمد الشركات على **التحليلات القائمة على البيانات** لتحسين الأداء واتخاذ القرارات. ومع ذلك، تواجه العديد من الشركات **التسريحات، وإعادة الهيكلة، أو قيود الميزانية**، مما يحد من قدرتها على الحفاظ على فرق تحليل بيانات متخصصة. يؤدي هذا إلى الحاجة إلى **حلول تحليل بيانات مؤتمتة** تسد هذه الفجوة وتضمن توفير رؤى موثوقة باستمرار.

## لمعالجة هذا التحدي، قمنا بتطوير **AutoInsight** وهو أداة ذكية **مؤتمتة لتحليل الأعمال** تعمل على تبسيط **معالجة البيانات وتحليلها**. تقوم الأداة **بأتمتة مرحلة المعالجة المسبقة للبيانات** وإجراء **ثلاثة أنواع من التحليل**: **التحليل التشخيصي** (لتحديد الاتجاهات الماضية والشذوذ)، **تحليل البيانات الاستكشافي (EDA)** (لكشف الأنماط والعلاقات)، و**التحليل التنبئي** (لتوقع الاتجاهات المستقبلية). يعتمد AutoInsight على **نموذج Facebook Prophet** لضمان **الكشف الدقيق عن الاتجاهات، وتحليل الموسمية، ورصد الشذوذ** دون الحاجة إلى تدخل يدوي.

## إلى جانب التنبؤ، يعزز AutoInsight تجربة المستخدم من خلال توفير **تقارير قابلة للتخصيص** تناسب احتياجات الشركات. كما يتضمن **روبوت دردشة مدعوماً بـ Gemini API**، مما يتيح للمستخدمين **التفاعل عبر النصوص والصور لفهم لوحات التحكم الخاصة بهم بعمق أكبر**. تساعد هذه الميزة المستخدمين في **تفسير البيانات، وفهم المؤشرات الرئيسية، واتخاذ قرارات مدروسة بكفاءة**.

## تظهر النتائج المستندة إلى بيانات حقيقية **دقة عالية في التنبؤ** وتحسين **اتخاذ القرارات القائمة على البيانات**. من خلال أتمتة التحليلات، يمكن للشركات **تحسين الأداء، ومراقبة الاتجاهات المالية، وتعزيز استراتيجيات التسويق، والمخزون، والتسعير**—مع تقليل الجهد اليدوي.

## ختامًا، يقدم **AutoInsight حلاً قابلاً للتطوير مدعوماً بالذكاء الاصطناعي**، مما يمكّن الشركات من الحفاظ على **قدرات تحليل البيانات الحيوية** بغض النظر عن قيود القوى العاملة. كما يساعد المؤسسات على اتخاذ **قرارات أكثر ذكاءً، وتحسين الكفاءة، والبقاء في الصدارة في سوق متغير باستمرار**.

**Table of Contents**

**Acknowledgements.…….……..……………..……………………………….i**

**Abstract………….………………...………………….………………………ii**

**List of Figures….……………………………….............…………….………iii**

**List of Tables…………...……………………………….……….……….…...iv**

**List of Abbreviations….……………………..…………………….…………v**

**List of Symbols…… ….……………………..…………………….…………vi**

**Chapter 1: Introduction………………………………………………...……1**

**1.1 Problem Definition…………………………………………1**

**1.2 Motivation…………………………………………………2**

**1.3 Objectives………………………………………………….2**

**1.4 Methodology** **………………………………………………3**

**1.5 Time plan………………………………………………….3**

**1.6 Thesis Outline ……………… ……………………… ……4**

**Chapter 2: Literature Review …………………….…..…………….5**

**Chapter 3: System Architecture and Methods ...…………………..**

**3.1 System Architecture ………………………………………**

**3.2 Description of methods and procedures used …………**

**Chapter 4: System Implementation and Results …..………………..**

**4.1 Dataset …..………………..………………..……………**

**4.2 Description of Software Tools Used ..………………..………**

**4.3 Stepup Configuration (hardware)**

**4.4 Experimental and Results …..…………………..……**

**Chapter 5: Run the Application....…………………………..…..**

**Chapter 6: Conclusion and Future Work………….…………………..**

**6.1 Conclusion…………………………………………………………**

**6.2 Future Work………………………………………………………**

**References……………………………………………………………….…….**

**List of Figures**

**For example**

Figure 1.1: Time Plane …..………………………………………………………11

Figure 1.2: System Architecture….………………………………………………20

**List of Tables**

**[State all tables in your documentation with page number]**

**Format**

Chapter number.Table number: Figure title……………………...………………….page number

**For example**

Table 1.1: Results of (….).……………...……………………………………21

Table 1.2: System Advantages and Disadvantages…….……………………………39

**List of Abbreviations**

EDA : **Exploratory Data Analysis**

API : **Application programming interface**

AI **: Artificial Intelligence**

ML : **Machine Learning**

Chapter One: Introduction

## In the modern business landscape, data analytics plays a critical role in driving decision-making and strategic planning.

## Recent studies indicate that over 90% of large companies utilize data analytics to enhance decision-making processes, leading to improved operational efficiency and profitability.

## According to McKinsey & Company, data-driven organizations are 23 times more likely to acquire customers, 6 times as likely to retain customers, and 19 times as likely to be profitable.

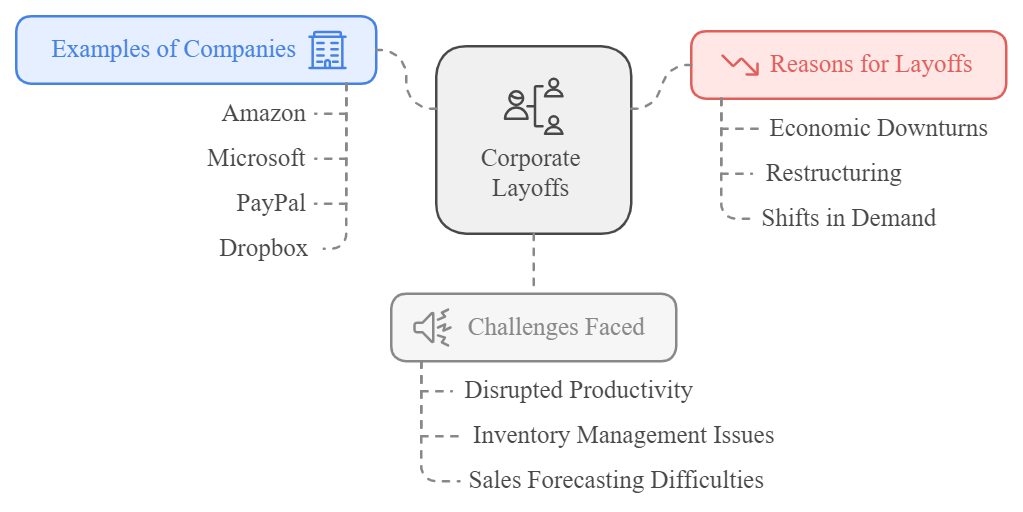
## The global big data and analytics market is projected to reach $684 billion by 2030, underscoring its significance in the corporate world.

## **1.1 Problem definition**

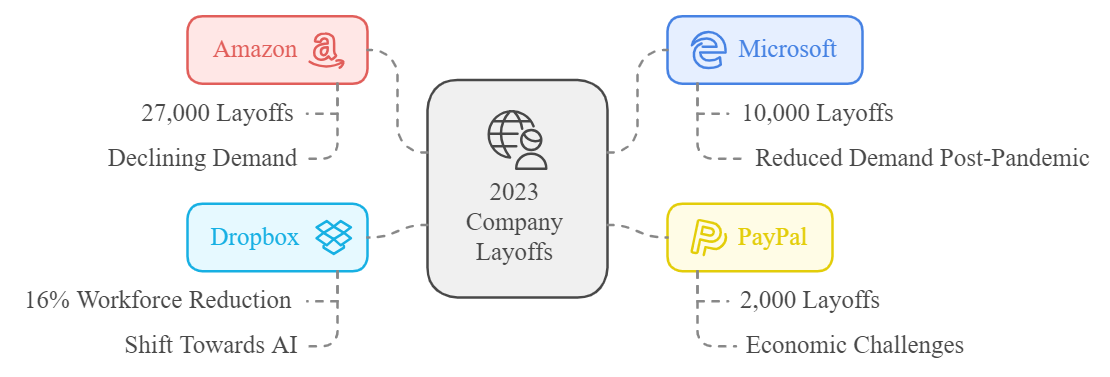
## Despite the undeniable benefits of data analytics, many companies face challenges related to layoffs, restructuring, and budget constraints, which limit their ability to maintain dedicated data analytics teams.

## As a result, there is an increasing demand for automated data analytics solutions that can mitigate these challenges by ensuring business continuity and informed decision-making without extensive human intervention.

## 



Layoff reasons Fig 1.1



Companies Layoffs Fig 1.2

## **1.2 Motivation**

The need for automated data analytics solutions stems from the following key challenges:

* **Large Companies:** Ensuring seamless data-driven decision-making during layoffs and restructuring phases.
* **Startups:** Supporting efficient data analysis amidst rapid growth or hiring limitations.
* **Data Analysts:** Enhancing productivity by automating repetitive tasks such as data cleaning and basic analysis, allowing focus on more complex insights.

Given these industry-wide challenges, this research aims to develop a comprehensive solution that empowers businesses to sustain and optimize their analytics capabilities, even in resource-constrained environments.

## **1.3 Objectives**

The primary objectives of this research are:

1. **Develop an Automated Data Analytics Platform:**
   * AutoInsight will provide tools for data analysis, forecasting, and machine learning models to enable businesses to make informed decisions efficiently.
2. **Ensure Business Continuity Amid Layoffs:**
   * The solution will assist organizations in maintaining productivity and decision-making capabilities despite workforce reductions.
3. **Enhance Cost Efficiency:**
   * By automating key data processes, businesses can reduce reliance on expensive analytics teams while sustaining operational efficiency.

## **1.4 Methodology**

This research employs various scientific methods to develop and validate the proposed solution:

1. **Automated Data Preprocessing:** Managing missing values, normalizing data, and structuring it for analysis.
2. **Diagnostic Analysis:** Identifying correlations, outliers, and factors influencing business performance.
3. **Exploratory Data Analysis (EDA):** Visualizing data trends, distributions, and relationships to uncover actionable insights.
4. **Predictive Analysis:** Implementing time series forecasting using the Prophet model to anticipate future business trends.
5. **Automated Report Generation:** Summarizing findings into customized reports to support strategic decision-making.
6. **Chatbot Integration (Gemini API):** Enabling interactive text and image-based conversations to assist users in navigating analytical dashboards.

By leveraging these methodologies, AutoInsight aims to enhance business intelligence, making data analytics more accessible and efficient for companies of all sizes.

## **1.5 Time Plan**

A structured timeline is established to ensure systematic development and implementation of the proposed solution. The project will be executed in phases, covering literature review, system design, implementation, testing, and evaluation.

**Chapter Two :** **Literature Review**

**2.1 Introduction**

The rapid advancement of data analytics and automation has transformed the way businesses operate.

The increasing reliance on data-driven decision-making underscores the necessity of automated analytics solutions, particularly during periods of workforce reductions and organizational restructuring.

This chapter provides a comprehensive review of the theoretical background and previous research related to automated data analytics, machine learning-driven insights, and decision-support systems.

By critically evaluating past studies and methodologies, this review highlights the existing gaps and justifies the need for an automated analytics platform like AutoInsight.

**2.2 Theoretical Background**

Data analytics is a multidisciplinary field that integrates various statistical, computational, and business intelligence techniques to extract meaningful insights from data.

The theoretical foundation of this project is based on several key concepts:

1. **Automated Data Analytics**
   * Automated analytics involves using artificial intelligence (AI) and machine learning (ML) to process and analyze data without significant human intervention.
   * It includes techniques such as data preprocessing, feature selection, and predictive modeling to enhance business decision-making.
2. **Machine Learning in Business Intelligence**
   * Predictive analytics models, such as time series forecasting (e.g., Prophet, ARIMA), help businesses anticipate trends and make proactive decisions.
   * Classification and clustering algorithms assist in identifying patterns and segmenting data for targeted analysis.
3. **Natural Language Processing (NLP) for Data Interpretation**
   * NLP enables interactive systems, such as chatbots (e.g., Gemini API), to interpret and respond to user queries, facilitating a seamless user experience in data analytics.
4. **Business Applications of AI-driven Analytics**
   * AI-driven analytics enhance operational efficiency by automating data processing, report generation, and strategic planning.
   * Companies use these solutions for customer insights, demand forecasting, and risk management.

By establishing a strong theoretical background, this research provides a structured foundation for the development of AutoInsight as a scalable and efficient analytics platform.

**2.3 Previous Studies and Related Works**

A critical review of previous studies highlights the evolution of data analytics automation and the impact of AI-driven decision-making tools:

1. **Automated Data Analytics in Industry**
   * Research by McKinsey & Company (2020) indicates that data-driven organizations are significantly more likely to achieve profitability and customer retention.
   * Studies have shown that automated analytics reduce operational costs while maintaining high accuracy in forecasting and reporting.
2. **Machine Learning for Business Forecasting**
   * A study by Hyndman & Athanasopoulos (2018) on time series forecasting demonstrated the effectiveness of statistical models like ARIMA and Prophet in predicting business trends.
   * Their work emphasizes the importance of accurate historical data and feature selection in improving model performance.
3. **AI-driven Decision Support Systems**
   * A review by Davenport & Harris (2017) on AI-powered business intelligence systems discusses the role of machine learning in automating insights for enterprises.
   * Their findings suggest that businesses adopting AI-based analytics see a marked improvement in operational agility and responsiveness.
4. **NLP and Chatbot Integration in Analytics**
   * Research by Budiu (2021) highlights the growing role of NLP in facilitating user interactions with analytics dashboards.
   * The integration of chatbot APIs, such as Gemini, allows non-technical users to navigate complex data insights seamlessly.

Through a critical evaluation of these studies, this research identifies key gaps in existing automated analytics solutions, particularly in the context of workforce reductions and budget constraints. AutoInsight aims to address these gaps by providing a cost-effective and scalable analytics platform.



Competitor Analysis Table 2.1

Chapter Three

REFERENCES

[1] H. Upadhyay, S. Shekhar, A. Vidyarthi, R. Prakash and R. Gowri, "Sales Prediction in the Retail Industry Using Machine Learning: A Case Study of BigMart,"  2023 pp. 1-6, doi: 10.1109/ELEXCOM58812.2023.10370313.7.

[2] S. Pang, "Retail Sales Forecast Based on Machine Learning Methods," 2022 pp. 357-361, doi: 10.1109/ICDSBA57203.2022.00030.

[3] A. Krishna, A. V, A. Aich and C. Hegde, "Sales-forecasting of Retail Stores using Machine Learning Techniques," 2018 pp. 160-166, doi: 10.1109/CSITSS.2018.8768765.

[4] M. Bilal, G. Ali, M. W. Iqbal, M. Anwar, M. S. A. Malik and R. A. Kadir, "Auto-Prep: Efficient and Automated Data Preprocessing Pipeline," in IEEE Access, vol. 10, pp. 107764-107784, 2022, doi: 10.1109/ACCESS.2022.3198662.

[5]<https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-data-driven-enterprise-of-2025>

[6]<https://www.keboola.com/blog/5-stats-that-show-how-data-driven-organizations-outperform-their-competition>

[7] <https://www.techmonitor.ai/digital-economy/big-tech/tech-layoffs-2023>

[8] <https://www.csv-master.com/en>

[9] <https://vizdium.com/>