Madanapalle Institute of Technology & Science Post Box No: 14, KadiriRoadAngallu (V), Madanapalle-517325, Annamayya District, Andhra Pradesh,India

PROJECT REPORT ON

CANCER AND DEATH DETECTION

Under the guidance of

Mr. AKASH V

CERTIFICATE OF INTERNSHIP

This is to certify that **P.Ayesha** bearing Registration Number:**24695A0501**, a student Madanapalle Institute of Technology & Science , KadiriRoad Angallu (V), Madanapalle-517325,Annamayya District, Andhra Pradesh,India has successfully completed an internship course from 16/06/2025 to 28/07/2025 at our organization.

STUDENT DECLARATION

I, **P.Ayesha**, Register Number: **24695A0501**, hereby declare that this report entitled **"Cancer and Death Detection"** using Machine Learning during the internship period from 16/06/2025 to 28/07/2025 at Prinston Smart Engineers, under the supervision and guidance of Mr. AKASH V...

Table of Contents:

EXECUTIVE SUMMARY

Brief Overview of the Internship Report

This project focuses on **Cancer and Death Detection using Machine Learning** techniques. The main objectives include identifying patients at high risk of death based on their cancer characteristics and medical records. The workflow involved:

- **Data Preprocessing**: Cleaning and encoding categorical features using tools like LabelEncoder. Missing values were handled appropriately to ensure model performance and reliability.
- **Logistic Regression**: A binary classification model was developed using logistic regression, which achieved a balanced accuracy. The model outputs whether a patient is at risk of death based on cancer-related input features.
- **Risk Score Calculation**: A custom "risk score" was developed to quantify the severity and predictability of outcomes, helping to visualize and interpret predictions.
- **Model Evaluation**: Accuracy, precision, recall, and F1-score were calculated. Confusion matrix and classification reports showed that the model had a reasonable balance between predicting deceased and living patients.
- **Visualization**: Histograms and distribution plots were used to analyze the relationship between risk score and survival status.

Overall, the project successfully demonstrates the application of machine learning in the healthcare domain, specifically for cancer prognosis and death prediction. It provides useful insights into how clinical data can be transformed into actionable intelligence.

Outcomes

The primary objective of this project was to develop a **machine learning model capable of predicting the likelihood of death in cancer patients** based on medical and demographic features. The outcomes of this project include:

- **Data Cleaning and Preprocessing**: Essential preprocessing steps such as handling missing values, encoding categorical variables, and normalizing features were performed to ensure data quality and consistency.
- **Exploratory Data Analysis (EDA)**: Visualizations and statistical analysis were used to understand the distribution of features and their correlation with the survival status.
- **Model Development**: A **Logistic Regression** model was implemented to classify patients based on their survival status. A **risk score** metric was also introduced to quantify the severity of each case.
- **Performance Evaluation**: The model was evaluated using metrics such as **accuracy**, **precision**, **recall**, and **F1-score**. These metrics confirmed the model's ability to differentiate between high-risk and low-risk patients.
- **Insights for Medical Decision Making**: The model and visualizations can support healthcare professionals in identifying patients requiring urgent intervention, thereby improving patient care and outcomes.

Conclusion

This project successfully demonstrates the use of **machine learning for cancer and death detection**, showcasing its potential in the **healthcare domain**. By using Logistic Regression and incorporating a risk score system, the project provided a practical framework for predicting patient outcomes based on clinical data.

The insights from this project can be valuable in supporting medical professionals in early diagnosis and personalized treatment planning. The approach highlights how data-driven tools can enhance clinical decision-making and contribute to better patient management.

2. INTRODUCTION

This internship focused on **cancer and death detection** using machine learning algorithms. The objective was to develop predictive models that could assess the likelihood of patient mortality based on cancer-related features. The dataset included both numerical and categorical variables such as patient demographics, cancer stage, treatment history, and other clinical indicators.

Techniques Used

- **Data Preprocessing**: The dataset required cleaning and preparation. Categorical variables such as 'Cancer Type', 'Hormone Therapy', 'Radiotherapy', 'Tumor Stage', 'Menopausal State', etc., were encoded using LabelEncoder to transform them into a numerical format suitable for model training.
- **Feature Engineering**: A new feature called **"Risk Score"** was introduced, which helped highlight the severity of patient health condition using a weighted combination of clinical variables.
- **Classification Models**:
    - **Logistic Regression**: This served as the baseline model. It provided a quick understanding of the relationship between variables and patient outcome.

- **Model Evaluation**:
    - **Accuracy**: Achieved up to **67.6%** and **97%** with Logistic Regression.
    - **Confusion Matrix**: Used to evaluate the true positives, true negatives, false positives, and false negatives.
    - **Classification Report**: Provided detailed metrics like precision, recall, and F1-score for both classes (Deceased and Living).

3.Description of the Organization

PRINSTON SMART ENGINEERS

"Imagination rules the world, we accomplish yours"

PRINSTON SMART ENGINEERS is a comprehensive Mechanical, Electrical, and Plumbing (MEP) Design consultancy, distinguished by its expertise in designing, constructing, installing, servicing, and upgrading Electro-Mechanical Systems & Networks, Utilities, and Equipment. their service offerings are extensive, covering every phase of a project from Engineering, Documentation, Submittals Approval, and Shop Drawings to Coordinating Drawings, Commissioning, Start-Up, and As-Built Documentation. they adhere meticulously to Contract Programs and Project Time Schedules, ensuring full compliance with Consultant Specifications. In addition to core engineering services, they provide a robust platform for engineering students to gain industry skills through live, instructor-led interactive online training sessions. their reach extends globally, with a significant presence in over 350 engineering colleges across India, delivering accessible and affordable learning solutions to learners.

Background:

PRINSTON SMART ENGINEERS was established in 2016 as a dedicated MEP Design consultancy. Recognizing the urgent need for skill development among engineering graduates in India, they swiftly expanded their services to include technical talks, training programs, and skill development initiatives for engineering students. India produces approximately 1.5 million engineering graduates each year, yet it is widely acknowledged that 75% of these graduates are unemployed. To bridge this gap, they have focused on providing essential skill training to engineering students, preparing them for the demands of the industry.

In 2018, the All India Council for Technical Education (AICTE) mandated internships for engineering students to ensure they gain exposure to industrial environments, current technologies relevant to their studies, and opportunities to enhance their technical and managerial skills. Responding to requests from numerous colleges, PRINSTON SMART ENGINEERS began offering high-quality internships led by their team of seasoned experts.

In 2020, they expanded their horizons through a strategic collaboration with Wedir-tech Trading Contracting & Services W.L.L in Doha, Qatar. This partnership was formed to enhance mutual benefits and improve service delivery. Wedir-tech Trading Contracting & Services W.L.L is a major player in the electro-mechanical field, known for its turnkey projects and a reputable customer base in the Middle East market.

Their Mission and Vision:

At PRINSTON SMART ENGINEERS, their mission is to deliver unparalleled engineering solutions and educational services that foster growth and innovation. they strive to set new standards in the MEP industry by consistently exceeding client expectations through their commitment to quality, safety, andsustainability. their vision is to be a global leader in MEP design and consultancy, known for our cutting- edge solutions, exceptional training programs, and dedication to creating a skilled workforce equipped to tackle the challenges of tomorrow.

1. **Skill Development and Training Programs:**

Recognizing the critical need for industry-ready professionals, PRINSTON SMART ENGINEERS

offers comprehensive training programs designed to equip engineering students with practical skills and knowledge. their training modules cover a wide range of topics, from fundamental engineering principles to advanced technical skills, ensuring that students are well-prepared for the demands of the industry. They offer live, instructor-led interactive online training sessions, providing students with hands-on experience and real-world insights.

1. **Strategic Partnerships and Collaborations:**

their collaboration with Wedir-tech Trading Contracting & Services W.L.L in Doha, Qatar, exemplifies their commitment to strategic partnerships that enhance capabilities and service offerings. This collaboration allows them to leverage Wedir-tech's extensive experience in the Middle Eastern market, providing their clients with comprehensive solutions that meet international standards.

1. **Future Plans and Growth:**

Looking ahead, PRINSTON SMART ENGINEERS is focused on expanding services and reach even further. They aim to establish a stronger presence in international markets, leveraging their expertise and strategic partnerships to deliver exceptional engineering solutions worldwide. their future plans also include the continuous development of training programs, ensuring that they remain at the forefront of industry education and skill development.

PRINSTON SMART ENGINEERS remains dedicated to core values of excellence, innovation, and sustainability. As they are committed to making a positive impact on the industry and the communities they serve, ensuring that they continue to be a trusted partner for all our engineering needs.

PRINSTON SMART ENGINEERS is unwavering in its dedication to delivering superior MEP solutions and educational services that promote growth and innovation. Their mission is to set new industry standards is driven by our commitment to quality,

safety, and sustainability, ensuring we consistently exceed client expectations. As they expand their global presence and refine their training programs, they remain focused on making a significant impact on the industry and the communities they serve.

their strategic collaborations, such as the partnership with Wedir-tech Trading Contracting & Services W.L.L in Doha, Qatar, enhance their capabilities and service offerings, allowing them to meet international standards and deliver comprehensive solutions. Future growth plans include broadening their services and reach, continuously improving their training programs to stay at the forefront of industry education.

With a strong foundation built on expertise, innovation, and a client-centric approach, PRINSTON SMART ENGINEERS is poised for continued success. they are committed to being a trusted partner for all engineering needs, delivering projects that are efficient, cost-effective, and sustainable. their dedication to continuous improvement and environmental preservation ensures they contribute positively to both industry advancement and societal well-being.

4.EXPERENTIAL LEARNING

Best Model Performance

- **Logistic Regression** gave the best balance in performance with an accuracy of **97% for cancer detection**
- **Logistic Regression** gave the best balance in performance with an accuracy of **67.6%**. for death detection
- The classification report revealed that the model was better at predicting the "Deceased" class, with a precision of **0.70** and recall of **0.80**, while predictions for the "Living" class were comparatively lower.

5.CONCLUSION

This project successfully demonstrated how machine learning can assist in **predicting cancer survival outcomes**. With an accuracy of 67.6%, the Logistic Regression model emerged as a dependable baseline. These results can aid medical professionals in risk assessment and further enhance early intervention strategies. Future improvements may include class balancing techniques, additional feature selection, or deep learning approaches for enhanced performance.

6.BIBLIOGRAPHY

Machine Learning Foundations:

- Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer.
- **Employee Performance Prediction**:
- Armstrong, M. B., & Landers, R. N. (2018). The Influence of Big Five Personality Traits and Shyness on Workplace Social Behaviors. Journal of Managerial Psychology, 33(3), 169–185.
- Barrick, M. R., & Mount, M. K. (1991). The Big Five Personality Dimensions and Job Performance: A Meta-Analysis. Personnel Psychology, 44(1), 1–26.
- **Employee Attrition Prediction**:
- Morrell, K., Loan-Clarke, J., & Wilkinson, A. (2004). Predicting job satisfaction: A new perspective on person–environment fit. Journal of Occupational and Organizational Psychology, 77(3), 321–338.

- Hasan, M. N., Bao, Q., Abbas, F., Zareei, A., & Tariq, U. (2020). Predicting Employee Attrition in Organizations Using Machine Learning Algorithms. Journal of Organizational and End User Computing, 32(2), 85–106.
- **Data Preprocessing and Feature Engineering**:
- Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. Journal of Machine Learning Research, 3, 1157–1182.
- Liu, H., & Motoda, H. (1998). Feature selection for knowledge discovery and data mining. Springer.
- **Machine Learning Algorithms**:
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning: with Applications in R. Springer.
- Müller, A. C., & Guido, S. (2016). Introduction to Machine Learning with Python: A Guide for Data Scientists. O'Reilly Media.
- **Model Evaluation and Validation**:
- Raschka, S., & Mirjalili, V. (2019). Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2. Packt Publishing Ltd.
- Kohavi, R., & Provost, F. (1998). Glossary of terms. Machine Learning, 30(2–3), 271–274.
- **Ethical Considerations in ML Projects**:
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389–399.
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. Big Data & Society, 3(2), 2053951716679679.
- **Interpretability and Explainability**:
- Lipton, Z. C. (2016). The Mythos of Model Interpretability. arXiv preprint arXiv:1606.03490.
- Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. In Proceedings of the 22nd ACM

SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 1135–1144).