

Name of Learners:- Shailaja Ramteke

Batch No:- B310

IT Vedant Institute of Bangalore

Internship Report Task 8

=====

Objective: Present the outcomes and take feedback for improvement.

Tasks: ○ Prepare a PowerPoint presentation summarizing the project workflow, insights, and results. ○ Discuss challenges and learning experiences during the internship. ○ Incorporate mentor feedback for further refinements.

Deliverables: ○ Final report and presentation. ○ A self-assessment document highlighting the internship learnings.

Week 8 Deliverables

Deliverables:-

1. PowerPoint Presentation :- Project Workflow: Slides detailing the project lifecycle, including problem definition, data preprocessing, model training, evaluation, and deployment. Visual representation of workflows, such as flowcharts, timelines, or process diagrams.

Key Insights and Results:- Summary of major findings and observations from exploratory data analysis (EDA) and model interpretation.

Performance metrics of the final model, including visualizations like confusion matrices, ROC curves, or performance comparison tables. Business impact of the results, explaining how the solution addresses the initial problem.

Challenges and Learnings: Slide(s) summarizing challenges faced during the project (e.g., data quality issues, algorithm selection, deployment hurdles). Key learning outcomes, including technical skills acquired and lessons from overcoming challenges.

2. Internship Reflection Document

A concise write-up discussing personal and professional growth during the internship.

Highlighted experiences with problem-solving, team collaboration, and applying technical skills in real-world scenarios.

3. Incorporated Mentor Feedback

Revised slides or content based on mentor feedback to enhance clarity, accuracy, and presentation quality.

Summary of feedback received and actions taken for refinement.

4. Final Presentation Ready for Delivery

A polished, well-structured presentation that includes:

Engaging visuals, concise content, and a logical flow.

Notes or talking points for each slide to guide the presenter.

5. Supporting Materials

Any supplementary documents, datasets, or resources referenced in the presentation.

Report

Report PPT link :- https://github.com/Shaila92/INTERNSHIP_TASK8

PowerPoint Presentation (Final Presentation Ready for Delivery)

Slide #	Title	Content
1	Title Slide	Project title, your name, internship organization, duration
2	Objective & Problem Statement	What is phishing? Why detect it? Project goal
3	Project Lifecycle Overview	Visual diagram: data → EDA → preprocessing → model → deployment
4	Dataset Summary	Key features, sample records, data shape
5	EDA Highlights	Example graphs: correlation heatmap, URL length distribution, class balance
6	Data Preprocessing	Missing value handling, scaling, encoding
7	Feature Engineering	New features created (e.g., special char ratios, URL length buckets)
8	Modeling Approaches	Algorithms tried (Logistic Regression, RF, Decision Tree, XGBoost)
9	Model Evaluation	Table: accuracy, precision, recall, F1, ROC-AUC
10	Performance Visuals	Confusion matrix, ROC curve, SHAP summary plot
11	Model Deployment	Architecture: User → API → Model → Response (Flask API diagram)
12	Key Insights	Most important features (SHAP), logical behavior confirmed
13	Business Impact	How model meets business objective (e.g. early phishing detection reduces risk)
14	Challenges Faced	Data imbalance, explainability, deployment setup
15	Learnings	Skills gained: ML, deployment, API creation, SHAP/LIME

Slide #	Title	Content
16	Mentor Feedback & Actions	Feedback received + improvements made
17	Future Work	Next steps (e.g., real-time detection, cloud scaling)
18	Thank You / Q&A	Contact info, acknowledgments

Slide 1: Title Slide

- Title: Phishing Detection Internship Final Presentation
- Subtitle: End-to-End Project Summary

Slide 2: Project Overview

- Objective: Develop, interpret, and deploy a machine learning model for phishing website detection.
- Scope: Data analysis, model building, explainability, deployment.
- Tech Stack: Python, scikit-learn, SHAP, Flask, Docker, AWS.

Slide 3: Project Workflow

- Diagram: User → Data → Preprocessing → Modeling → Evaluation → Explainability → Deployment → User Access
- Timeline or flowchart illustrating each phase.

Slide 4: Exploratory Data Analysis (EDA)

- **Key observations from data:**
 - Features like length_url, nb_dots, and domain_age were strong indicators.
 - Class imbalance handled.
- **Visuals:**
 - Heatmap of feature correlation
 - Distribution of length_url

Slide 5: Data Preprocessing & Feature Engineering

- Handled missing values, duplicates.
- Feature selection: correlation + importance.

- New features: length-based, special characters.
- Data split: 80% train / 20% test.

Slide 6: Model Training & Evaluation

- Models: Logistic Regression, Decision Tree, Random Forest.
- Final model: Random Forest (best accuracy & generalization).
- Metrics table: Accuracy, Precision, Recall, F1, ROC-AUC.
- ROC curve visual.

Slide 7: Model Explainability

- SHAP summary plot showing key features.
- SHAP dependency plot example.
- Insight: length_url, nb_dots, domain_age drive decisions.

Slide 8: Deployment

- Local Flask API + Docker container for portability.
- Cloud: AWS EC2 instance hosting the model API.
- Architecture: User → API → Model → Response
- Screenshot of working API endpoint (if available).

Slide 9: Challenges & Learnings

- **Challenges:**
 - Managing high-dimensional data.
 - Cloud deployment configurations.
- **Learnings:**
 - Enhanced Python, ML, Flask, Docker, AWS skills.
 - Improved problem-solving, debugging.

Slide 10: Mentor Feedback & Refinements

- Feedback: Improve clarity of visualizations, simplify API endpoints.
- Actions: Enhanced visuals, cleaned API code.

Slide 11: Conclusion & Business Impact

- Delivered a reliable, interpretable model.
- Solution helps in early detection of phishing websites.
- Framework is scalable for future enhancements.

Internship Reflection Document Summary

Introduction

The internship focused on building, explaining, and deploying a phishing website detection model using machine learning.

Key Learnings

- Applied ML techniques: EDA, feature engineering, model selection.
- Built APIs with Flask and containerized with Docker.
- Gained hands-on experience with SHAP for explainability.

Personal & Professional Growth

- Strengthened technical skills: Python, ML, deployment.
- Gained confidence in real-world problem-solving.
- Enhanced communication through reporting and presentation.

Challenges

- Data quality (imbalanced data, noisy features).
- Deployment setup (port conflicts, dependency issues).
- Balancing accuracy and interpretability.

Feedback and Improvements

- Incorporated mentor feedback: enhanced SHAP plots, improved documentation.
- Added security checks to API routes.

Personal Growth

- Improved coding practices, model explainability understanding.
- Stronger confidence in deploying ML models.

Future Aspirations

- Explore cloud-native solutions and real-time detection systems.

Supporting Materials

- Final Jupyter notebook (clean + commented)
- app.py (Flask API)
- Dockerfile
- Sample input/output files
- Visual assets (plots, diagrams)

Key Experiences

- Hands-on model building and deployment.
- Worked on feedback cycles for improvement.

Teamwork

- Collaborated with mentor, accepted guidance for refining deliverables.
-