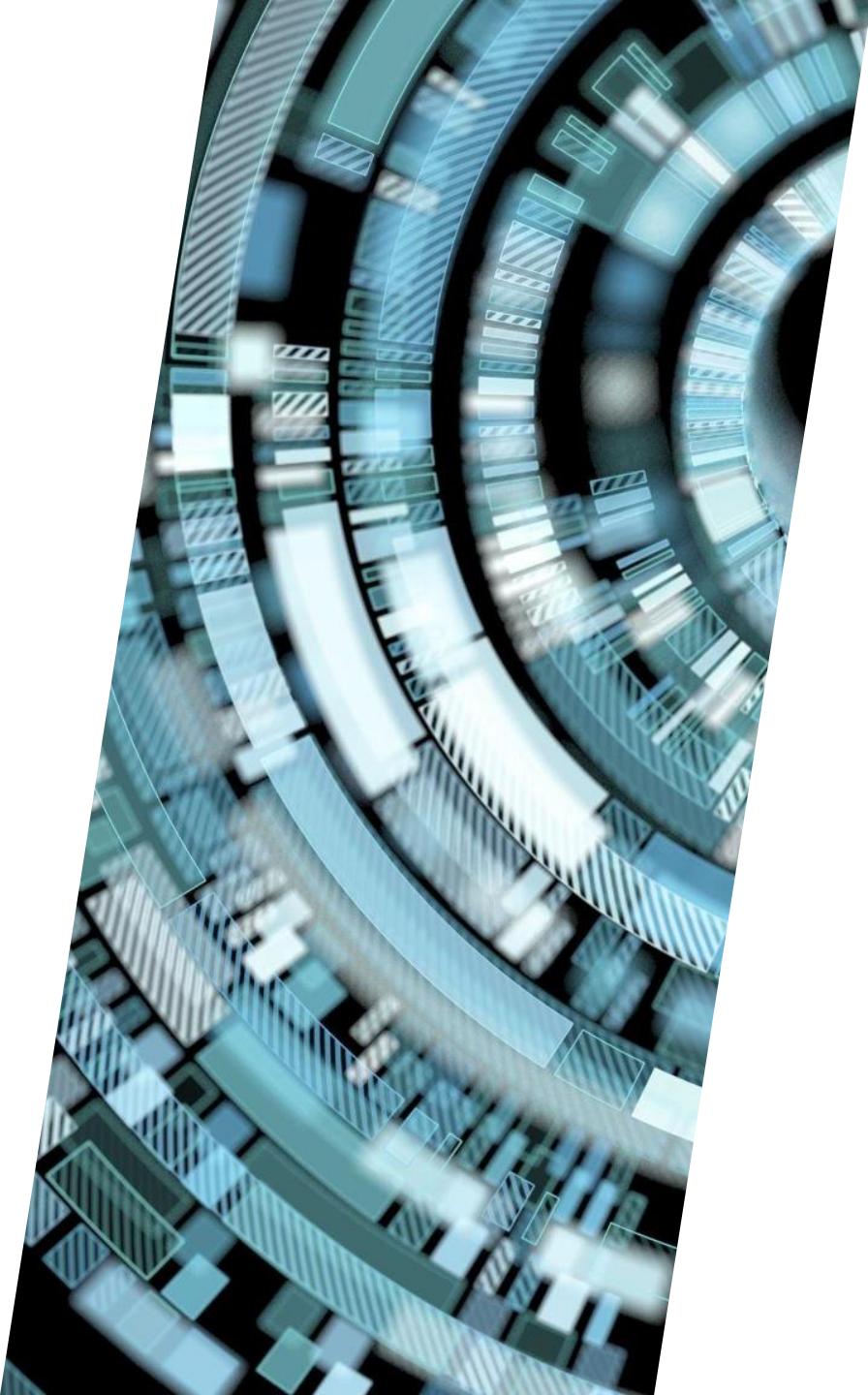# COMPARATIVE ANALYSIS OF SUPPORT VECTOR MACHINE PERFORMANCE ACROSS GEOPHYSICS, FINANCE, AND MEDICAL DIAGNOSTICS

AGENDA & PURPOSE OF STUDY

- Introduction & Methodology

- Comparative Findings (3 Domains)

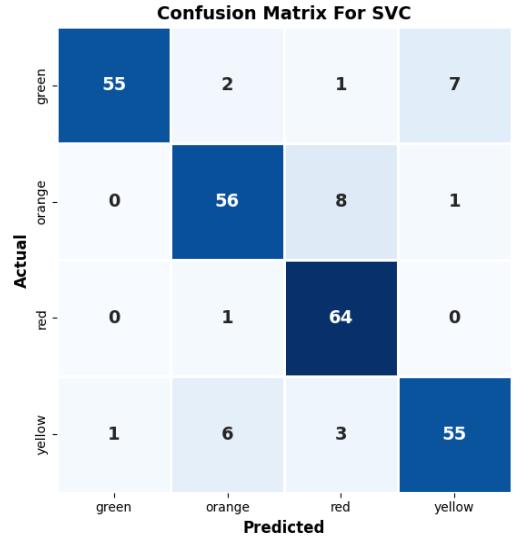- Discussion & Model Limitations

INTRODUCTION & METHODOLOGY

1. *Purpose:* To evaluate the SVM's performance against three distinct, real-world problems (Medical, Financial, Geophysical)

2. *Data Cleaning:* Data was first cleaned by addressing duplicate records and structural inconsistencies

3. *Data Preparation:* Missing values – mode imputation, outliers – managed via class-wise median imputation. Standardization to all numerical values

4. *Validation:* Data partitioned with an 80%-20% ratio. Stratified K-Fold Cross-Validation – for imbalanced Data

5. *Model Setup:* SVM (RBF kernel) was optimized by tuning the regularization parameter to achieve the best performance for each domain
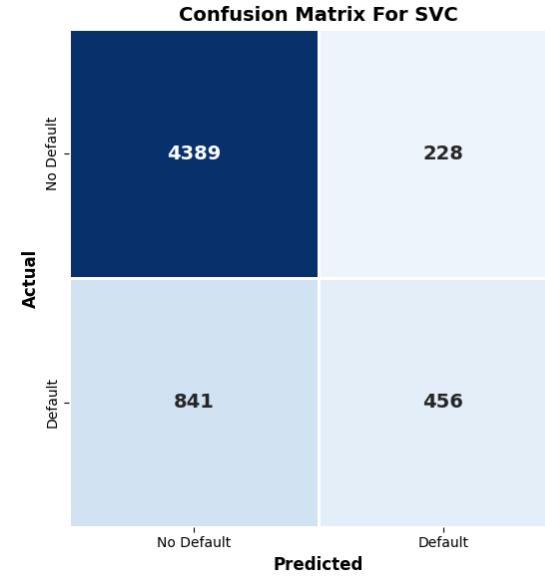
COMPARATIVE FINDINGS

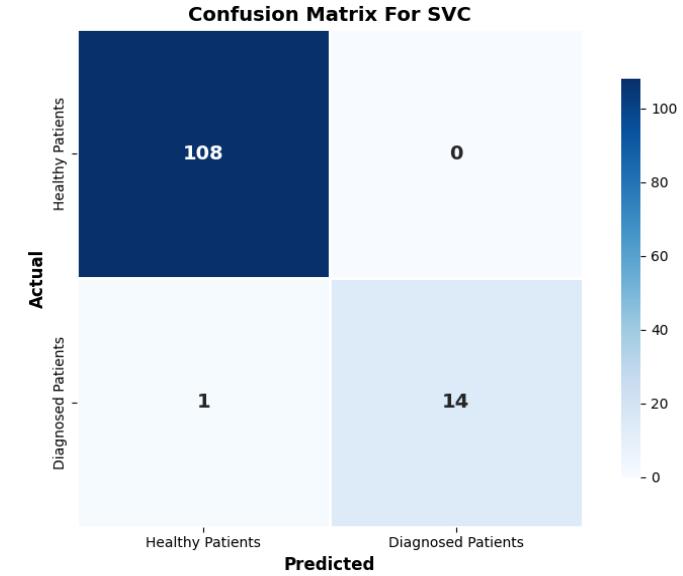| DATASET	ACCURACY	PRECISION	RECALL	F1-Score
HCV Data	0.9918	0.9919	0.9918	0.9917
Credit Card Default Data	0.8192	0.3515	0.6666	0.4603
Earthquake Alert Data	0.8846	0.8904	0.8846	0.8847

CONFUSION MATRICES

Earthquake Alert Prediction

Default of Credit Card Clients

HCV Data (Hepatitis C Virus)

DISCUSSION & MODEL LIMITATIONS

Model Strengths

- High capability when data features were clean and separable

Model Limitations

- Weakness: high recall to severe class imbalance and overlapping features

Analytical Insight

- Low Recall 0.66 confirms that Accuracy is an inappropriate metric when the model struggles with the minority class.