Title: Building a Decision Tree Workflow

Subtitle: A Step-by-Step Approach to Model Building, Evaluation, and Insights

ISSUE / PROBLEM

The goal of this project is to develop a Decision Tree model to make accurate predictions on a given dataset. Decision Trees are a popular choice due to their interpretability and minimal preprocessing requirements. The challenge is to build an effective model that balances predictive performance while avoiding overfitting.

IMPACT

- Improved decision-making by levering a model that provides transparent and explainable predictions.
- Reduction in preprocessing efforts compared to other models like Neural Networks or Support Vector Machines.
- Identification of key features driving predictions, aiding domain experts in refining their decision-making strategies.
- This project demonstrates the effectiveness of Decision Trees in predictive modeling and highlights best practices for achieving robust performance while mitigating overfitting risks.

RESPONSE

This project follows a structured workflow consisting of:

1. **Data Exploration:** Understanding dataset characteristics and distributions.
2. **Model Building:** Splitting the dataset into training and testing sets.
3. **Evaluation:** Assessing model performance using metrics like accuracy, precision, recall, and F1 score.
4. **Visualization:** Plotting data to analyze features.
5. **Considerations:** Exploring potential overfitting issues and hyperparameter tuning strategies.

KEY INSIGHTS

- **Performance Metrics:** This model achieved an **accuracy of 93.5%**, with a **precision of 94.3%**, **recall of 93.9%**, and **F1 score of 94.1%** indicating strong predictive power.
- **Confusion Matrix Analysis:** A high proportion of **true positives and true negatives**, with relatively low false positives and false negatives.
- **Overfitting Risk:** Decision Trees can be prone to overfitting, making **hyperparameter tuning essential** to optimize model generalization.