



Data Science Project Lifecycle & Tools

Prasad Deshmukh

1. Problem Definition and Planning:

- Define the problem statement and project goals.

TOOLS

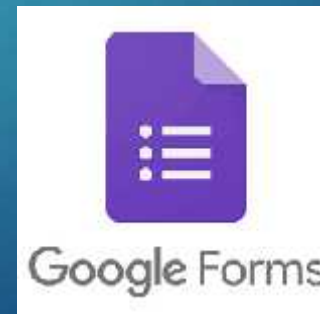


Prasad Deshmukh

2. Data Acquisition

- Identify relevant data sources and gather the required data.

TOOLS



Prasad Deshmukh

3. Data Cleaning and Preprocessing

- Clean the data by handling missing values, outliers, and inconsistencies.

TOOLS



Prasad Deshmukh

4. Exploratory Data Analysis (EDA)

- ▶ Understand the data through visualization and statistical analysis.
- ▶ Identify patterns, correlations, and outliers.

TOOLS



SciPy

Prasad Deshmukh

5. Model Selection and Development

- ▶ Select appropriate machine learning models based on the problem and data characteristics.
- ▶ Train and evaluate models using suitable metrics.

TOOLS

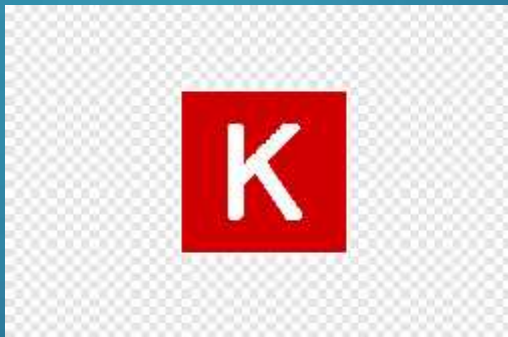


Prasad Deshmukh

6. Model Tuning and Optimization

- ▶ Fine-tune model hyperparameters using techniques like grid search or random search.
- ▶ Optimize models for improved performance and generalization.

TOOLS



Prasad Deshmukh

7. Model Evaluation and Validation

- ▶ Evaluate models using appropriate validation techniques such as cross-validation or train-test splits.
- ▶ Measure model performance using metrics specific to the problem.

TOOLS



Prasad Deshmukh

8. Deployment and Integration

- ▶ Deploy the model into a production environment.
- ▶ Integrate the model with existing systems or applications.

TOOLS



Prasad Deshmukh

9. Monitoring and Maintenance

- ▶ Continuously monitor model performance and data quality.
- ▶ Retrain or update models as needed.

TOOLS



Prasad Deshmukh

10. Documentation and Communication

- ▶ Document the entire project, including data sources, methodology, and findings.
- ▶ Communicate results to stakeholders through reports or presentations.

TOOLS



Prasad Deshmukh



It's important to note that the specific tools used may vary depending on individual preferences, project requirements, and the data science ecosystem at the time. The list above provides a general overview of the stages and commonly used tools in a data science project life cycle.

Prasad Deshmukh



THANK YOU

Prasad Deshmukh