A Complete 20–30 Page Guide to Matplotlib for Machine Learning

Prepared for ML Learners

November 30, 2025

Contents

Chapter 1

Introduction to Matplotlib

1.1 What is Matplotlib?

Matplotlib is the foundational visualization library for Python. It is used directly or indirectly in nearly every machine learning workflow. Libraries such as Seaborn, Scikit-Learn, and Yellowbrick internally depend on Matplotlib.

1.2 Why Learn Matplotlib for ML?

Machine learning requires:

- Data exploration
- Feature distribution analysis
- Outlier detection
- Training curve visualization
- Model comparison
- Image visualization for computer vision

Matplotlib enables all these tasks with high control and flexibility.

1.3 Installing Matplotlib

```
pip install matplotlib
```

1.4 Two APIs: Pyplot vs Object-Oriented

Matplotlib provides:

- **Pyplot API** — simple, MATLAB-style
- **Object-Oriented API** — preferred for ML work

Chapter 2

Matplotlib Basics

2.1 Creating Basic Plots

2.1.1 Line Plot

```
import matplotlib.pyplot as plt

x = [1,2,3,4,5]
y = [1,4,9,16,25]

plt.plot(x, y)
plt.title("Basic Line Plot")
plt.xlabel("x-values")
plt.ylabel("y-values")
plt.show()
```

2.2 Scatter Plot

Useful for regression, patterns, and clusters.

2.3 Bar Plot

Used in classification tasks and feature importance.

Chapter 3

Understanding the Object-Oriented API

3.1 Why the OO-API Matters in ML

- More control
- Clean multi-plot figures
- Easier to integrate with ML frameworks (TensorFlow/PyTorch)

3.2 Creating Figures and Axes

```
fig, ax = plt.subplots()
ax.plot([1,2,3],[1,4,9])
ax.set_title("OO Plot")
plt.show()
```

Chapter 4

Subplots and Multi-Panel Figures

4.1 Using subplots

```
fig, axes = plt.subplots(2, 2, figsize=(10,6))

axes[0,0].plot(...)
axes[0,1].scatter(...)
axes[1,0].bar(...)
axes[1,1].hist(...)

plt.tight_layout()
plt.show()
```

This is used heavily in ML:

- Training vs validation curves
- Original vs augmented images
- Multi-feature distributions

Chapter 5

Styling and Customization

5.1 Colors, Markers, Line Styles

```
plt.plot(x, y, color='red', linestyle='--', marker='o')
```

5.2 Legends, Grid, Figure Size

5.3 Matplotlib Style Sheets

```
plt.style.use('ggplot')
```

Chapter 6

Histograms and Distribution Visualization

6.1 Histograms

Essential for:

- identifying skewness
- outliers
- data spread

6.2 Kernel Density Plots

(Using Seaborn but Matplotlib-compatible)

6.3 Box Plots and Violin Plots

Chapter 7

Scatter Plots in Machine Learning

7.1 Color-based Labels

```
plt.scatter(x, y, c=labels)
```

7.2 Size Variations

7.3 Colormap Visualization

Chapter 8

Heatmaps for ML

8.1 Correlation Matrix

```
plt.imshow(corr_matrix, cmap='viridis')
plt.colorbar()
```

8.2 Confusion Matrix (Manual and Sklearn)

Chapter 9

Machine Learning-Specific Visualizations

9.1 Training and Validation Curves

9.1.1 Loss Curve

```
plt.plot(history['loss'])
plt.plot(history['val_loss'])
plt.legend(['train', 'validation'])
plt.title("Loss Curve")
```

9.2 Accuracy Curve

9.3 ROC and PR Curves

9.4 Feature Importance Plots

9.5 Clustering Visualization (k-means)

Chapter 10

Matplotlib for Computer Vision

10.1 Displaying Images

```
plt.imshow(image, cmap='gray')
plt.axis('off')
```

10.2 Displaying Multiple Images

10.3 Visualizing Feature Maps

10.4 Overlaying Bounding Boxes

Chapter 11

Saving and Exporting Figures

11.1 Saving with DPI

```
plt.savefig("plot.png", dpi=300)
```

11.2 Saving as PDF, SVG

Chapter 12

Advanced Techniques and Tips

12.1 Interactive Plots with Jupyter

12.2 Animations with Matplotlib

12.3 3D Plots

12.4 Embedding Plots into ML Reports

Chapter 13

Full ML Visualization Workflows

13.1 Case Study: Regression Problem

- Scatter plots
- Residual plots
- Predicted vs actual

13.2 Case Study: Classification Problem

- Confusion matrix
- ROC curve
- Feature importance

13.3 Case Study: CNN Training

- Loss curves
- Image previews
- Augmentation visualization

Chapter 14

Summary and Best Practices

14.1 What You Should Focus On Most

- Line plots (loss/accuracy)
- Histograms
- Scatter plots
- Heatmaps
- Image visualization
- Subplots (comparisons)

14.2 Final Notes

Mastering these techniques prepares you for all ML visualization tasks from basic EDA to advanced deep learning workflows.