ML Community 2-Month Summer Roadmap

Duration: June 15 - August 15, 2025
Daily Commitment: 1 hour (Monday-Friday), 1 hour Kaggle (Saturday), Sunday OFF

Week 1: Python Fundamentals & NumPy Mastery

June 16-21

Daily Schedule:

- **Monday (June 16):** Python essentials - variables, data types, control structures, functions, loops (1 hour)
- **Tuesday (June 17):** Python advanced - list comprehensions, lambda functions, decorators, error handling (1 hour)
- **Wednesday (June 18):** NumPy fundamentals - array creation, indexing, slicing, reshaping (1 hour)
- **Thursday (June 19):** NumPy operations - mathematical operations, broadcasting, boolean indexing (1 hour)
- **Friday (June 20):** NumPy advanced - linear algebra, statistics, random, array manipulation functions (1 hour)
- **Saturday (June 21): KAGGLE DAY** - Explore Kaggle platform, choose a beginner dataset, basic EDA with NumPy (1 hour)

Weekly Coverage:

Python: Variables, data types, control flow, functions, modules, OOP basics, file I/O, exception handling, comprehensions **NumPy:** Array creation, indexing, slicing, mathematical operations, broadcasting, linear algebra, statistics, random number generation

---

Week 2: Pandas Complete & Data Visualization Mastery

June 23-28

Daily Schedule:

- **Monday (June 23):** Pandas fundamentals - Series, DataFrames, data loading (CSV, Excel, JSON), basic operations (1 hour)
- **Tuesday (June 24):** Pandas data manipulation - indexing, selection, filtering, sorting, groupby, aggregations (1 hour)
- **Wednesday (June 25):** Pandas advanced - merging, joining, pivot tables, reshaping, time series basics (1 hour)

- **Thursday (June 26):** Data cleaning mastery - missing values, duplicates, data types, string operations, outliers (1 hour)
- **Friday (June 27):** Visualization complete - Matplotlib (plots, subplots, styling), Seaborn (statistical plots, themes) (1 hour)
- **Saturday (June 28): KAGGLE DAY** - Apply Pandas skills to Kaggle dataset, complete data cleaning and visualization (1 hour)

Weekly Coverage:

Pandas: DataFrames, Series, data loading, indexing, filtering, groupby, merging, pivot tables, time series, string methods, data cleaning **Visualization:** Matplotlib (line, bar, scatter, histogram, subplots), Seaborn (boxplots, heatmaps, pair plots, distribution plots)

Week 3: Machine Learning Foundations

June 30 - July 5

Daily Schedule:

- **Monday (June 30):** ML fundamentals - supervised vs unsupervised, regression vs classification, ML workflow (1 hour)
- **Tuesday (July 1):** Scikit-learn ecosystem - datasets, preprocessing, model selection, metrics (1 hour)
- **Wednesday (July 2):** Data preprocessing - scaling, normalization, encoding categorical variables, feature selection (1 hour)
- **Thursday (July 3):** Linear regression - theory, implementation, evaluation metrics (MSE, R^2) (1 hour)
- **Friday (July 4):** Model validation - train/test split, cross-validation, overfitting/underfitting (1 hour)
- **Saturday (July 5): KAGGLE DAY** - Implement linear regression on Kaggle dataset, submission practice (1 hour)

Weekly Goal:

Master ML fundamentals and linear regression with proper validation techniques.

Week 4: Classification Algorithms

July 7-12

Daily Schedule:

- **Monday (July 7):** Logistic regression - theory, implementation, probability interpretation (1 hour)
- **Tuesday (July 8):** Decision trees - theory, implementation, feature importance, pruning (1 hour)
- **Wednesday (July 9):** Random forests - ensemble methods, bagging, feature importance (1 hour)

- **Thursday (July 10):** Support Vector Machines - theory, kernels, parameter tuning (1 hour)
- **Friday (July 11):** Classification metrics - accuracy, precision, recall, F1-score, confusion matrix, ROC-AUC (1 hour)
- **Saturday (July 12): KAGGLE DAY** - Classification competition, compare multiple algorithms (1 hour)

Weekly Goal:

Master classification algorithms and evaluation metrics.

Week 5: Advanced ML & Unsupervised Learning

July 14-19

Daily Schedule:

- **Monday (July 14):** K-Nearest Neighbors - theory, implementation, distance metrics, curse of dimensionality (1 hour)
- **Tuesday (July 15):** Naive Bayes - theory, different types, text classification applications (1 hour)
- **Wednesday (July 16):** Clustering - K-means, hierarchical clustering, DBSCAN, evaluation metrics (1 hour)
- **Thursday (July 17):** Dimensionality reduction - PCA, t-SNE, feature selection techniques (1 hour)
- **Friday (July 18):** Hyperparameter tuning - GridSearchCV, RandomizedSearchCV, pipeline optimization (1 hour)
- **Saturday (July 19): KAGGLE DAY** - Apply clustering or dimensionality reduction to dataset (1 hour)

Weekly Goal:

Complete traditional ML algorithms and unsupervised learning techniques.

Week 6: Ensemble Methods & Advanced Techniques

July 21-26

Daily Schedule:

- **Monday (July 21):** Ensemble methods - bagging, boosting, stacking concepts (1 hour)
- **Tuesday (July 22):** Gradient boosting - XGBoost, LightGBM, CatBoost implementation (1 hour)
- **Wednesday (July 23):** Feature engineering - creating features, polynomial features, interactions (1 hour)
- **Thursday (July 24):** Model interpretation - feature importance, SHAP values, LIME (1 hour)
- **Friday (July 25):** Advanced preprocessing - handling imbalanced data, SMOTE, feature scaling strategies (1 hour)

- **Saturday (July 26): KAGGLE DAY** - Implement ensemble methods, feature engineering on competition (1 hour)

Weekly Goal:

Master ensemble methods and advanced ML techniques.

Week 7: Deep Learning Fundamentals

July 28 - August 2

Daily Schedule:

- **Monday (July 28):** Neural networks theory - perceptrons, backpropagation, activation functions (1 hour)
- **Tuesday (July 29):** TensorFlow/Keras setup - building first neural network, compilation, training (1 hour)
- **Wednesday (July 30):** Deep learning for tabular data - architecture design, regularization, dropout (1 hour)
- **Thursday (July 31):** Convolutional Neural Networks - CNN architecture, convolution, pooling layers (1 hour)
- **Friday (August 1):** CNN implementation - image classification, data augmentation, transfer learning basics (1 hour)
- **Saturday (August 2): KAGGLE DAY** - Deep learning approach to current competition or image dataset (1 hour)

Weekly Goal:

Understand neural networks and implement CNN for image tasks.

Week 8: Advanced Deep Learning & Specialization

August 4-9

Daily Schedule:

- **Monday (August 4):** Recurrent Neural Networks - RNNs, LSTMs, GRUs for sequence data (1 hour)
- **Tuesday (August 5):** Natural Language Processing - text preprocessing, embeddings, sentiment analysis (1 hour)
- **Wednesday (August 6):** Transfer learning - pre-trained models, fine-tuning, feature extraction (1 hour)
- **Thursday (August 7):** Model optimization - learning rate scheduling, batch normalization, advanced optimizers (1 hour)

- **Friday (August 8):** MLOps basics - model saving, loading, deployment concepts, monitoring (1 hour)
- **Saturday (August 9): KAGGLE DAY** - Apply advanced techniques to competition, ensemble deep learning models (1 hour)

Weekly Goal:

Advanced deep learning techniques and deployment preparation.

Week 9: Final Projects & Portfolio Development

August 11-15

Daily Schedule:

- **Monday (August 11):** Individual project planning - choose dataset, define problem, project structure (1 hour)
- **Tuesday (August 12):** Project development - data preprocessing, EDA, baseline model (1 hour)
- **Wednesday (August 13):** Project development - advanced modeling, hyperparameter tuning, evaluation (1 hour)
- **Thursday (August 14):** Project finalization - documentation, visualization, presentation preparation (1 hour)
- **Friday (August 15):** Project presentations - showcase work, peer feedback, portfolio discussion (1 hour)

Weekly Goal:

Complete end-to-end ML project for portfolio.

Saturday Kaggle Schedule (Every Week):

Week 1-2: Foundation Building

- Explore datasets, practice data manipulation
- Focus on EDA and visualization skills

Week 3-4: ML Application

- Implement learned algorithms on competitions
- Focus on model building and evaluation

Week 5-6: Advanced Techniques

- Apply ensemble methods and advanced preprocessing
- Compete seriously in ongoing competitions

Week 7-8: Deep Learning Integration

- Use neural networks where appropriate
- Experiment with different architectures

Week 9: Final Push

- Complete competition submissions
- Document learnings and strategies

Deep-ML.com Integration:

- **Week 1:** NumPy and Python problems (3-4 problems)
- **Week 2:** Pandas problems (3-4 problems)
- **Week 3-4:** Basic ML algorithm problems (2-3 problems each week)
- **Week 5-6:** Advanced ML problems (2-3 problems each week)
- **Week 7-8:** Deep learning problems (2-3 problems each week)

LeetCode Pandas Problems:

- **Week 2:** Focus on Pandas LeetCode problems (5-6 problems)
- **Ongoing:** 1-2 problems weekly for practice

Resources & Tools:

Week 1-2: Python, NumPy, Pandas, Matplotlib, Seaborn, Jupyter **Week 3-6:** Scikit-learn, XGBoost, LightGBM, SHAP **Week 7-8:** TensorFlow/Keras, OpenCV (if needed) **Week 9:** Streamlit/Flask for deployment (optional)

Success Metrics:

- Complete comprehensive understanding of Python, NumPy, Pandas (Weeks 1-2)
- Implement all major ML algorithms (Weeks 3-6)
- Build neural networks and CNNs (Weeks 7-8)
- Active Kaggle participation with improving scores
- Complete end-to-end project with deployment
- Solve 15-20 deep-ml.com problems total
- Master 10+ Pandas LeetCode problems

This intensive roadmap ensures complete mastery of foundations before diving into ML, with consistent Kaggle practice throughout!