## **Roadmap Overview: From 0 to 100 in Machine Learning**

### **📍 Phase 1: Foundations (Beginner – Month 1)**

**Objective:** Build your math, stats, Python, and ML basics.

* 📘 **Module 1: Math & Stats for ML**
  + Linear Algebra (vectors, matrices, dot product, eigenvalues)
  + Probability (conditional, Bayes theorem, distributions)
  + Statistics (mean, median, mode, variance, skewness, hypothesis testing)
  + Calculus (partial derivatives, gradients – needed for optimization)
* 🐍 **Module 2: Python Essentials**
  + Numpy, Pandas
  + Matplotlib, Seaborn
  + Writing functions, classes, debugging, vectorized operations
* 🧠 **Module 3: Machine Learning Basics**
  + What is ML? Types of ML: Supervised, Unsupervised, Reinforcement
  + Train/test split, Overfitting/Underfitting
  + Performance metrics (Accuracy, Precision, Recall, F1, AUC)
  + First models: Linear Regression, Logistic Regression, KNN

### **📍 Phase 2: Intermediate (Month 2–3)**

**Objective:** Cover classical algorithms and model building.

* 🌲 **Module 4: Supervised Learning**
  + Decision Trees, Random Forests
  + Gradient Boosting (XGBoost, LightGBM)
  + Support Vector Machines (SVM)
  + Regularization (L1, L2)
* 📊 **Module 5: Model Evaluation & Tuning**
  + Cross-validation (K-Fold, Stratified K-Fold)
  + Hyperparameter tuning (GridSearchCV, RandomizedSearchCV, Optuna)
  + Feature engineering & selection
* 🧪 **Module 6: Unsupervised Learning**
  + Clustering: K-Means, DBSCAN, Hierarchical
  + Dimensionality Reduction: PCA, t-SNE, UMAP
  + Anomaly Detection: Isolation Forests, One-Class SVM

### **📍 Phase 3: Advanced ML (Month 4–5)**

**Objective:** Master high-level concepts, deployable systems, and problem-solving.

* 🔮 **Module 7: Time Series & NLP**
  + Time Series Forecasting: ARIMA, Prophet, LSTMs
  + NLP: TF-IDF, Word2Vec, Transformers (BERT), Hugging Face
* 🧩 **Module 8: Deep Learning (with PyTorch/TensorFlow)**
  + Neural Networks: MLPs, CNNs, RNNs
  + Backpropagation, Activation Functions, BatchNorm, Dropout
  + Advanced models: Autoencoders, GANs, Attention, Transformers
* 🏭 **Module 9: End-to-End ML Projects**
  + Full project pipelines: EDA → Model → Evaluation → Deployment
  + ML system design: scalable systems, latency, retraining strategies

### **📍 Phase 4: Interview & Industry Prep (Ongoing)**

**Objective:** Get hired into a top ML/DS role.

* 🔥 **Module 10: Interview-Level Case Studies**
  + Fraud Detection, Recommendation Systems, Search Ranking
  + Feature stores, real-time inference, online/offline evaluation
* 🧠 **Module 11: ML System Design Interviews**
  + ML architecture, batch vs streaming, monitoring
  + Questions like: "Design a spam detection system", "A/B test search relevance"
* ❓ **Module 12: Top Interview Questions**
  + Why XGBoost over Random Forest?
  + Bias-variance tradeoff?
  + What if your model is 99% accurate but failing in production?
  + ML fairness, explainability, interpretability