ZERO TO ADVANCED

ML→DL→Transformers→LLMs → MLOps→Projects →Interview Prep.

FINAL AI/ML ROADMAP (WITH FULL MATH & STATS)

This is your complete journey from zero → advanced → Google-level.

Daily requirement: 3–4 hours minimum

FINAL AI/ML ROADMAP (WITH FULL MATH & STATS)
This is your complete journey from *zero → advanced → Google-level*.

STAGE 1 — Foundational Math + Python

Goal: Build the mathematical and programming base required for all ML.

A. Mathematics (Master these topics early)

1. Linear Algebra

- Vectors, matrices
- Matrix multiplication
- Dot product, cross product
- Rank, determinant, inverse
- Eigenvalues, eigenvectors
- Orthogonality, projections
- Singular Value Decomposition (SVD)

2. Calculus

- Limits, continuity
- Derivatives
- Partial derivatives
- Gradients
- Chain rule
- Jacobians
- Taylor series
- Integrals
- Gradient descent intuition (very important)

3. Probability & Statistics

- Random variables
- Mean, median, variance, standard deviation
- Expectation
- Probability distributions (Normal, Bernoulli, Binomial, Poisson, Uniform)
- Bayes theorem
- Maximum likelihood estimation
- Hypothesis testing (Z-test, t-test)
- Confidence intervals

4. Discrete Math (Used for ML engineering)

- Sets
- Functions
- Graph theory basics
- Combinatorics
- Logic
- Recurrence relations

B. Python Foundations

- Python syntax, loops, functions
- OOP in Python
- NumPy (vectorized operations)
- Pandas (data manipulation)
- Matplotlib + Seaborn (visualization)
- Basic debugging
- Writing clean, modular code

STAGE 2 — Core Machine Learning

Goal: Build strong ML knowledge with full mathematical depth.

A. ML Theory

- Supervised vs unsupervised learning
- Train–validation–test splits
- Bias–variance tradeoff
- Underfitting & overfitting
- Cross-validation
- Regularization (L1, L2)
- Loss functions (MSE, MAE, cross-entropy)

B. ML Algorithms + Math Behind Them

Linear Models

- Linear regression

    - Normal equation
    - Gradient descent optimization
- Logistic regression

Tree-Based Models

- Decision trees
- Random forest
- Gradient boosting
- XGBoost/LightGBM/CatBoost (important for interviews)

Clustering

- K-Means (distance, centroids, variance)
- Hierarchical clustering
- Gaussian Mixture Models (GMM)
- EM algorithm (Expectation Maximization)

Dimensionality Reduction

- PCA (Singular values, eigenvectors)
- LDA

C. ML Engineering Skills

- Data preprocessing
- Feature engineering
- Handling missing data
- Scaling/normalization
- One-hot encoding
- Handling imbalanced data (SMOTE, class weights)
- Model evaluation metrics:
    - Accuracy
    - Precision
    - Recall
    - F1-score
    - ROC-AUC
    - Confusion matrix

STAGE 3 — Deep Learning Foundations

Goal: Learn neural networks with full mathematical intuition.

A. Deep Learning Math

- Tensors
- Forward propagation
- Backpropagation
- Activation functions (ReLU, tanh, sigmoid, softmax)
- Gradient descent variants (SGD, Adam, RMSProp)
- Vanishing/exploding gradients
- Loss functions: cross-entropy, hinge loss

B. Neural Networks

- Perceptron
- Multilayer perceptrons
- Feed-forward networks
- Hyperparameter tuning
- Batch normalization
- Dropout
- Optimizers

C. Tools

- PyTorch (preferred for AI researchers)
- TensorFlow/Keras (easy for production)

STAGE 4 — Advanced Deep Learning

Goal: Move into the real AI world—CV, NLP, LLMs.

A. Computer Vision

- Convolutions
- CNNs
- ResNet, DenseNet
- Transfer learning
- Data augmentation
- Object detection basics (YOLO, SSD)
- Vision Transformers (ViT)

B. Natural Language Processing

- Text preprocessing
- Tokenization (Byte Pair Encoding, WordPiece)
- Embeddings (Word2Vec, GloVe)
- RNN, GRU, LSTM
- Attention mechanism
- Transformers
- BERT
- GPT architecture
- Large Language Models (LLMs) training pipeline

C. Generative Models

- Variational Autoencoders (VAE)
- GANs (DCGAN, StyleGAN)
- Diffusion Models (Stable Diffusion basic concepts)

STAGE 5 — ML Engineering & Production

Goal: Become employable for top companies.

A. MLOps

- Data pipelines
- Model versioning
- CI/CD for ML
- Docker
- Kubernetes
- MLFlow
- Weights & Biases
- Automated model retraining workflows
- GPU/TPU training pipelines

B. System Design for ML

- Large-scale data processing
- Distributed training (PyTorch Distributed, Horovod)
- Serving models in production (gRPC, FastAPI)
- Caching, load balancing
- Real-time inference design

STAGE 6 — Research-Level AI Knowledge

Goal: Reach Google AI researcher-level depth.

A. Read and implement research papers

Start with:

- Attention Is All You Need
- BERT
- YOLOv8
- CLIP
- Stable Diffusion
- AlphaFold

B. Implement custom models from scratch

- Implement backprop manually
- Build a transformer manually
- Build your own LLM fine-tuning pipeline
- Train small models on custom datasets

STAGE 7 — Portfolio + Interviews

Goal: Become hire-ready.

A. Portfolio projects

Build at least 6 high-quality projects, including:

- End-to-end ML system
- Deep learning CV model
- NLP transformer-based model
- LLM fine-tuning project
- AI product (full stack ML system)
- Research paper reproduction

B. Interview Prep

- DSA (Google-level: Trees, DP, Graphs)
- ML theory interview prep
- System design
- ML system design
- Behavioural preparation

- ***Mathematics & Statistics***
- ***ML theory***
- ***Deep learning***
- ***Research-level AI***
- ***ML engineering***
- ***Google-level system design***
- ***Portfolio-building***
- ***Interview readiness***

- **Advance DAA & DSA in JAVA**