ToolTechSavvy

AI Engineer Roadmap

Complete Mastery Checklist

Phase 1: Mathematical Foundations

Linear Algebra

- ■ Vectors and vector spaces
- ■ Matrices and matrix operations
- ■ Eigenvalues and eigenvectors
- ■ Singular Value Decomposition (SVD)
- ■ Matrix factorization techniques
- ■ Linear transformations
- ■ Orthogonality and projections

Calculus

- ■ Differential calculus (derivatives, chain rule)
- ■ Partial derivatives
- ■ Gradient and directional derivatives
- ■ Integral calculus
- ■ Multivariable calculus
- ■ Optimization fundamentals
- ■ Taylor series and approximations

Probability & Statistics

- ■ Probability theory basics
- ■ Conditional probability and Bayes' theorem

- Random variables and distributions
- Expected value and variance
- Common distributions (Normal, Bernoulli, Poisson, etc.)
- Statistical inference
- Hypothesis testing
- Confidence intervals
- Maximum Likelihood Estimation (MLE)
- Bayesian statistics
- Sampling methods

Optimization

- Gradient descent and variants
- Stochastic Gradient Descent (SGD)
- Convex optimization
- Constrained optimization
- Lagrange multipliers
- Newton's method
- Coordinate descent

Phase 2: Programming Foundations

Python Programming

- Python syntax and data structures
- Object-oriented programming (OOP)
- Functional programming concepts
- File I/O operations
- Exception handling
- Decorators and generators
- Context managers

- Multithreading and multiprocessing
- Asynchronous programming (async/await)

Data Structures & Algorithms

- Arrays and linked lists
- Stacks and queues
- Trees and graphs
- Hash tables and dictionaries
- Sorting algorithms
- Search algorithms
- Dynamic programming
- Recursion
- Time and space complexity (Big O)

Software Engineering

- Git and version control
- Code documentation
- Unit testing
- Integration testing
- CI/CD pipelines
- Code review practices
- SOLID principles
- Clean code practices
- Agile methodologies

Phase 3: Data Science & Analysis

NumPy

- Array creation and manipulation

- Broadcasting
- Indexing and slicing
- Mathematical operations
- Linear algebra with NumPy
- Random number generation
- Memory optimization

Pandas

- DataFrames and Series
- Data loading (CSV, Excel, JSON, SQL)
- Data cleaning and preprocessing
- Handling missing values
- Data filtering and selection
- Grouping and aggregation
- Merging and joining datasets
- Time series analysis
- Data transformation

Data Visualization

- Matplotlib basics and advanced plots
- Seaborn for statistical visualization
- Plotly for interactive visualizations
- Dashboard creation (Streamlit, Dash)
- Data storytelling principles

Exploratory Data Analysis

- Descriptive statistics
- Data distribution analysis
- Correlation analysis

- Feature relationships
- Outlier detection
- Data profiling

Phase 4: Machine Learning Fundamentals

Supervised Learning

- Linear regression
- Logistic regression
- Decision trees
- Random forests
- Gradient boosting (XGBoost, LightGBM, CatBoost)
- Support Vector Machines (SVM)
- K-Nearest Neighbors (KNN)
- Naive Bayes classifiers

Unsupervised Learning

- K-means clustering
- Hierarchical clustering
- DBSCAN
- Principal Component Analysis (PCA)
- t-SNE
- UMAP
- Autoencoders
- Anomaly detection methods

Model Evaluation & Selection

- Train/validation/test splits
- Cross-validation techniques

- Metrics (accuracy, precision, recall, F1)
- ROC curves and AUC
- Confusion matrices
- Bias-variance tradeoff
- Overfitting and underfitting
- Regularization (L1, L2, Elastic Net)
- Hyperparameter tuning
- Grid search and random search

Feature Engineering

- Feature scaling and normalization
- Encoding categorical variables
- Feature creation and extraction
- Feature selection techniques
- Dimensionality reduction
- Handling imbalanced datasets

ML Libraries

- Scikit-learn mastery
- XGBoost implementation
- LightGBM usage
- CatBoost for categorical data

Phase 5: Deep Learning Fundamentals

Neural Network Basics

- Perceptrons and activation functions
- Feedforward neural networks
- Backpropagation algorithm
- Loss functions

- Optimization algorithms (SGD, Adam, RMSprop)
- Weight initialization strategies
- Batch normalization
- Dropout and regularization
- Learning rate scheduling

Deep Learning Frameworks

- PyTorch fundamentals
- TensorFlow/Keras fundamentals
- Computational graphs
- Automatic differentiation
- GPU acceleration basics
- Model saving and loading
- Transfer learning implementation

Convolutional Neural Networks (CNNs)

- Convolution operations
- Pooling layers
- CNN architectures (LeNet, AlexNet, VGG)
- ResNet and skip connections
- Inception networks
- MobileNet and EfficientNet
- Image classification
- Object detection (YOLO, R-CNN, SSD)
- Image segmentation (U-Net, Mask R-CNN)
- Data augmentation techniques

Recurrent Neural Networks (RNNs)

- Basic RNN architecture
- Vanishing and exploding gradients

- Long Short-Term Memory (LSTM)
- Gated Recurrent Units (GRU)
- Bidirectional RNNs
- Sequence-to-sequence models
- Time series forecasting
- Text generation

Phase 6: Natural Language Processing (NLP)

NLP Fundamentals

- Text preprocessing (tokenization, stemming, lemmatization)
- Bag of Words (BoW)
- TF-IDF
- Word embeddings (Word2Vec, GloVe)
- Named Entity Recognition (NER)
- Part-of-speech tagging
- Sentiment analysis

Transformer Architecture

- Attention mechanisms
- Self-attention and multi-head attention
- Positional encoding
- Transformer architecture (encoder-decoder)
- BERT and variants
- GPT architecture
- T5 and other models

Large Language Models (LLMs)

- Pre-training vs fine-tuning
- Transfer learning in NLP

- Prompt engineering
- Few-shot and zero-shot learning
- In-context learning
- Instruction tuning
- RLHF (Reinforcement Learning from Human Feedback)
- Model quantization
- LoRA and parameter-efficient fine-tuning

NLP Libraries

- NLTK
- spaCy
- Hugging Face Transformers
- LangChain
- OpenAI API integration
- Vector databases (Pinecone, Weaviate, ChromaDB)

Advanced NLP Applications

- Question answering systems
- Text summarization
- Machine translation
- Chatbot development
- Information retrieval
- Retrieval-Augmented Generation (RAG)

Phase 7: Computer Vision

Image Processing

- ■ Image filtering and convolution
- ■ Edge detection
- ■ Image transformations
- ■ Color space conversions
- ■ Histogram equalization
- ■ OpenCV mastery

Advanced Computer Vision

- ■ Image classification at scale
- ■ Multi-object detection and tracking
- ■ Instance segmentation
- ■ Semantic segmentation
- ■ Panoptic segmentation
- ■ Pose estimation
- ■ Facial recognition systems
- ■ OCR (Optical Character Recognition)
- ■ Image generation (GANs, Diffusion models)
- ■ Style transfer
- ■ Super-resolution

Video Analysis

- ■ Video classification
- ■ Action recognition
- ■ Video object tracking
- ■ Motion estimation
- ■ 3D reconstruction

Phase 8: Advanced AI Topics

Generative AI

- Generative Adversarial Networks (GANs)
- Variational Autoencoders (VAEs)
- Diffusion models (DDPM, Stable Diffusion)
- Text-to-image generation
- Image-to-image translation
- Audio generation
- Video generation

Reinforcement Learning

- Markov Decision Processes (MDP)
- Q-Learning
- Deep Q-Networks (DQN)
- Policy gradients
- Actor-Critic methods
- Proximal Policy Optimization (PPO)
- Monte Carlo Tree Search (MCTS)
- Multi-agent RL
- Inverse RL

Graph Neural Networks

- Graph theory basics
- Graph Convolutional Networks (GCN)
- Graph Attention Networks (GAT)
- GraphSAGE

- Message passing frameworks
- Knowledge graphs

Multimodal AI

- Vision-language models (CLIP, ALIGN)
- Image captioning
- Visual question answering
- Audio-visual learning
- Cross-modal retrieval

Phase 9: MLOps & Production

Model Deployment

- REST API development (Flask, FastAPI)
- Model serialization (pickle, ONNX)
- Docker containerization
- Kubernetes basics
- Model versioning
- A/B testing frameworks
- Blue-green deployments
- Canary releases

Cloud Platforms

- AWS (SageMaker, EC2, S3, Lambda)
- Google Cloud Platform (Vertex AI, Cloud Run)
- Azure ML
- Serverless architectures
- Cloud storage solutions

MLOps Tools

- MLflow for experiment tracking
- Weights & Biases (wandb)
- DVC (Data Version Control)
- Airflow for workflow orchestration
- Kubeflow
- TensorFlow Serving
- TorchServe
- BentoML

Monitoring & Maintenance

- Model performance monitoring
- Data drift detection
- Concept drift detection
- Logging and alerting
- Model retraining pipelines
- Feature store implementation
- Model explainability (SHAP, LIME)

Performance Optimization

- Model compression
- Pruning techniques
- Knowledge distillation
- Quantization (INT8, FP16)
- TensorRT optimization
- ONNX Runtime
- Batch processing optimization

- Caching strategies

Phase 10: Specialized Domains

Time Series Analysis

- ARIMA models
- Prophet
- LSTM for time series
- Temporal Convolutional Networks
- Anomaly detection in time series
- Forecasting techniques

Recommender Systems

- Collaborative filtering
- Content-based filtering
- Matrix factorization
- Neural collaborative filtering
- Deep learning for recommendations
- Cold start problem solutions
- Evaluation metrics for recommender systems

Audio Processing

- Speech recognition (ASR)
- Text-to-speech (TTS)
- Audio classification
- Music generation
- Speaker recognition

- Audio enhancement
- Librosa library

Healthcare AI

- Medical image analysis
- Disease prediction models
- Drug discovery applications
- Clinical NLP
- HIPAA compliance
- FDA regulations for AI/ML

Finance & Trading

- Algorithmic trading strategies
- Portfolio optimization
- Risk assessment models
- Fraud detection
- Credit scoring
- Market prediction

Phase 11: Ethics & Responsible AI

AI Ethics

- Bias in AI systems
- Fairness metrics and mitigation
- Privacy-preserving machine learning
- Differential privacy
- Federated learning

- Explainable AI (XAI)
- Interpretability techniques
- AI safety principles
- Adversarial robustness
- AI alignment

Legal & Compliance

- GDPR compliance
- Data protection regulations
- Model governance
- Intellectual property in AI
- Liability considerations
- Ethical guidelines and frameworks

Phase 12: Research & Innovation

Research Skills

- Reading and understanding research papers
- Implementing papers from scratch
- Experimental design
- Statistical significance testing
- Ablation studies
- Benchmarking methodologies
- Writing technical documentation
- Conference paper writing

Staying Current

- ■ Following top AI conferences (NeurIPS, ICML, CVPR, ACL)
- ■ ArXiv paper reading habits
- ■ AI research blogs and newsletters
- ■ Contributing to open-source projects
- ■ Building a portfolio
- ■ Technical blog writing
- ■ Kaggle competitions
- ■ Personal projects

Phase 13: Soft Skills & Career

Communication

- ■ Technical presentation skills
- ■ Stakeholder management
- ■ Data storytelling
- ■ Documentation writing
- ■ Team collaboration
- ■ Code review practices

Project Management

- ■ Scoping AI projects
- ■ Estimating timelines and resources
- ■ Risk assessment
- ■ Stakeholder expectation management
- ■ Proof of concept (PoC) development
- ■ MVP strategy

Business Acumen

- Understanding business metrics
- ROI calculation for AI projects
- Problem identification
- Solution architecture
- Cost-benefit analysis
- Product thinking

Learning Resources

Books

- "Deep Learning" by Ian Goodfellow, Yoshua Bengio, Aaron Courville
- "Hands-On Machine Learning" by Aurélien Géron
- "Pattern Recognition and Machine Learning" by Christopher Bishop
- "Natural Language Processing with Transformers" by Lewis Tunstall et al.
- "Designing Data-Intensive Applications" by Martin Kleppmann

Online Courses

- Andrew Ng's Machine Learning (Coursera)
- Deep Learning Specialization (Coursera)
- Fast.ai courses
- Stanford CS229, CS230, CS224N, CS231N
- Full Stack Deep Learning

Practice Platforms

- Kaggle competitions and datasets
- LeetCode for algorithms

- HackerRank for problem solving
- GitHub for portfolio building
- Personal blog for documentation

Milestones & Projects

Beginner Projects

- Linear regression from scratch
- Image classifier with CNN
- Sentiment analysis with LSTM
- Recommendation system

Intermediate Projects

- Object detection application
- Chatbot with transformers
- Time series forecasting system
- End-to-end ML pipeline with MLflow

Advanced Projects

- Fine-tune LLM for specific domain
- Build RAG system
- Deploy production ML system
- Contribute to major open-source AI project
- Implement latest research paper
- Build multimodal AI application

Total Topics: 400+

Estimated Timeline: 12-24 months of dedicated study

Remember: Master fundamentals before moving to advanced topics. Practice consistently, build projects, and contribute to the community!

Track your progress by checking off items as you master them. Good luck on your AI engineering journey!