

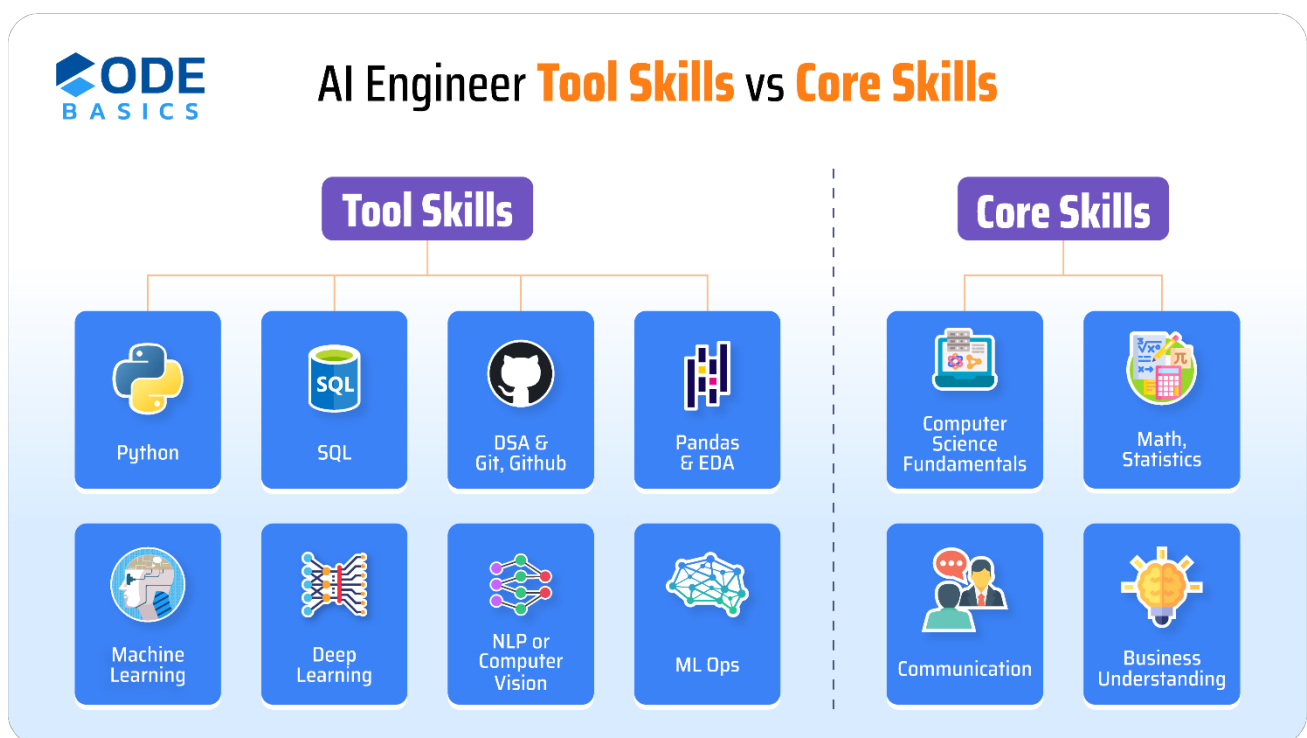
# AI Engineer Roadmap for Beginners

Following is the roadmap to learning **AI Engineer** (also known as **ML Engineer**) skills for a total beginner. It includes FREE learning resources for technical skills (or tool skills) and soft (or core) skills 🏃

**Prerequisites:** You must have skills or interests to build skills in Coding and Math. Without these two you cannot become an AI engineer.

Total Duration: **8 Months** (4 hours of study Every Day)

Also, **AI Engineer = Data Scientist + Software Engineer**



## Week 0: Do Proper Research and protect yourself from SCAMS.

Unfortunately, a lot of systematic scams are happening in ed tech, especially in the data field where aspirants are provided with false promises like a 100% job guarantee or trapped into "Masterclasses" which are nothing but sales pitches to upsell their low-grade courses at exorbitant prices. You need to do complete research about the market and mentors before starting your journey. Providing you the links to a few posts that we have made in this regard which will support your research.

Even though these posts are **NOT** sufficient, do your additional research.

- <https://bit.ly/4at9Jaw>
- <https://bit.ly/477lOOs>
- <https://bit.ly/3GPD7dp>

## Week 1 and 2: Computer Science Fundamentals

---

- **Topics**
  - Data representation: Bits and Bytes, Storing text and numbers, Binary number system.
  - Basics of computer networks, IP addresses, Internet routing protocol
  - UDP, TCP, HTTP, and The World Wide Web
  - Programming basics: variables, strings, and numbers, if condition, loops
  - Algorithm basics
- **Learning Resources**
  - Khan Academy course: <https://bit.ly/42DUXtW>
  - In the above course, only follow the first 4 sections (1) Digital Information (2) The Internet (3) Programming (4) Algorithms. Completing the remaining sections is optional. Do it if you have time and interest.

## Week 3 and 4: Beginners Python

---

- **Topics**
  - Variables, Numbers, Strings
  - Lists, Dictionaries, Sets, Tuples
  - If condition, for loop
  - Functions, Lambda Functions
  - Modules (pip install)
  - Read, Write files
  - Exception handling
  - Classes, Objects
- **Learning Resources**
  - Track A (Free)
    - Python Tutorials (Codebasics) on YouTube (first 16 videos)  
- <https://bit.ly/3X6CCC7>
    - Corey's Python Tutorials: <https://bit.ly/3uqUgaZ>
    - Codebasics python HINDI tutorials  
- <https://bit.ly/3vmXrgw>

- Track B (Affordable Fees)
  - Python course: <https://codebasics.io/courses/python-for-beginner-and-intermediate-learners>
- **LinkedIn - Core Skill**
  - ☐ Create a professional-looking LinkedIn profile.
    - Have a clear profile picture and banner image.
    - Add tags such as: Open to work etc.
  - ☐ Use this LinkedIn Checklist to create a profile: [Click here.](#)
- **Assignment**
  - ☐ Track A: Finish all these exercises: <https://bit.ly/3k1mof5>
  - ☐ Track B: Finish exercises and quizzes for relevant topics
  - ☐ Create a professional-looking LinkedIn profile.

## Week 5 and 6: Data Structures and Algorithms in Python

---

- **Topics**
  - Data structures basics, Big O notation
  - Data structures: Arrays, Linked List, Hash Table, Stack, Queue
  - Data structures: Tree, Graph
  - Algorithms: Binary search, Bubble sort, quick sort, merge sort
  - Recursion
- **Learning Resources**
  - DSA YouTube Playlist: <https://bit.ly/3uiW2Lf>
- **Motivation**
  - How Kaggle helped this person become ML engineer: <https://bit.ly/3RFVruy>
- **Assignment**
  - ☐ Finish all these exercises in this same playlist: <https://bit.ly/3uiW2Lf>

## Week 7, 8: Advance Python

---

- **Topics**

- Inheritance, Generators, Iterators
- List Comprehensions, Decorators
- Multithreading, Multiprocessing

- **Learning Resources**

- Python Tutorials (Codebasics) on YouTube (17<sup>th</sup> to 27<sup>th</sup> video)  
- <https://bit.ly/3X6CCC7>

- **Assignment**

- ☐ Finish all these exercises in this same playlist: <https://bit.ly/3X6CCC7>

- **Core/Soft Skills**

- **Linkedin**

- Start following prominent AI influencers.
  - Daliana Liu: <https://www.linkedin.com/in/dalianaliu/>
  - Nitin Aggarwal: <https://www.linkedin.com/in/ntnaggarwal/>
  - Steve Nouri: <https://www.linkedin.com/in/stevenouri/>
  - Dhaval Patel: <https://www.linkedin.com/in/dhavalays/>
- Increase engagement.
  - Start commenting meaningfully on AI and career-related posts.
  - Helps network with others working in the industry build connections.
  - Learning and brainstorming opportunity.
- Remember ***online presence is a new form of resume***

- **Business Fundamentals - Soft Skill**

- Learn business concepts from ThinkSchool and other YT Case Studies
- Example: How Amul beat competition: <https://youtu.be/nnwqtZiYMxQ>

- **Discord**

- Start asking questions and get help from the community. This post shows how to ask questions the right way: <https://bit.ly/3I70Ebl>
- Join codebasics discord server: <https://discord.gg/r42Kbuk>

- **Assignment**

- ☐ Write meaningful comments on at least **10 AI related LinkedIn posts**
- ☐ Note down your key learnings from **3 case studies** on ThinkSchool and share them with your friend.

## Week 9: Version Control (Git, Github)

---

- **Topics**

- What is the version control system? What is Git and GitHub?
- Basic commands: add, commit, push.
- Branches, reverting change, HEAD, Diff and Merge
- Pull requests.

- **Learning Resources**

- YT playlist (codebasics): <https://bit.ly/3SECQQ7>
- YT playlist (Corey): <https://bit.ly/3T0Yrmb>

- **Motivation**

- Mechanical to Deep Learning Engineer: <https://bit.ly/48IX9aR>

- **Core/Soft Skills**

- Presentation skills
  - Death by PowerPoint: <https://youtu.be/lwpi1Lm6dFo>

## Week 10, 11: SQL

---

- **Topics**

- Basics of relational databases.
- Basic Queries: SELECT, WHERE LIKE, DISTINCT, BETWEEN, GROUP BY, ORDER BY
- Advanced Queries: CTE, Subqueries, Window Functions
- Joins: Left, Right, Inner, Full
- Database creation, indexes, stored procedures.

- **Learning Resources**

- Track A
  - Khan academy SQL course: <https://bit.ly/3WFku20>
  - <https://www.w3schools.com/sql/>
  - <https://sqlbolt.com/>
  - YT video: <https://youtu.be/Rm0xH2Vpfi0?si=6ZLK8A5LvGqN4NmT>
- Track B
  - SQL course for data professionals: <https://codebasics.io/courses/sql-beginner-to-advanced-for-data-professionals>

- **Assignment**

- ☐ Participate in SQL resume project challenge on <https://codebasics.io/>
  - Link: <https://codebasics.io/challenge/codebasics-resume-project-challenge/7>
  - These challenges help you improve technical skills, soft skills and business understanding.
- ☐ Make a LinkedIn post with a submission of your resume project challenge  
Sample post: <https://bit.ly/48Bg5mB>

## Week 12: Numpy, Pandas, Data Visualization

---

- **Tech Skills**

- **Numpy**
  - numpy YouTube playlist: <https://bit.ly/3GTppa8>
- **Pandas, Matplotlib, Seaborn**
  - Go through chapter 3 in this course (entire chapter is free):  
<https://codebasics.io/courses/math-and-statistics-for-data-science>

## Week 13, 14, 15, 16: Math & Statistics for AI

---

- **Math and Statistics for AI**

- Topics to Learn
  - Basics: Descriptive vs inferential statistics, continuous vs discrete data, nominal vs ordinal data
  - Linear Algebra: Vectors, Metrics, Eigenvalues and Eigenvectors
  - Calculus: Basics of integral and differential calculus

- Basic plots: Histograms, pie charts, bar charts, scatter plot etc.
- Measures of central tendency: mean, median, mode
- Measures of dispersion: variance, standard deviation
- Probability basics
- Distributions: Normal distribution
- Correlation and covariance
- Central limit theorem
- Hypothesis testing: p value, confidence interval, type 1 vs type 2 error, Z test
- Learning Resources
  - Track A (Free)
    - Learn the above topics from this excellent Khan academy course on statistics and probability.
    - Course link: <https://www.khanacademy.org/math/statistics-probability>
    - While doing khan academy course, when you have doubts, use statquest YouTube channel: <https://www.youtube.com/@statquest>
    - Use this free YouTube playlist: <https://bit.ly/3QrSXis>
    - Another great youtube channel: <https://www.youtube.com/@3blue1brown>
  - Track B (Affordable Fees)
    - Learn the key concepts of Math and Statistics that lay the foundations for a strong data science career: <https://codebasics.io/courses/math-and-statistics-for-data-science>
- **Assignment**
  - ☐ Finish all exercises in this playlist: <https://bit.ly/3QrSXis>
  - ☐ Finish all exercises in Khan academy course.
  - ☐ Track B: Finish exercises and quizzes for relevant topics.

## Week 17: Exploratory Data Analysis (EDA) 🧐

---

- **Exploratory Data Analysis (EDA)**

- <https://www.kaggle.com/code?searchQuery=exploratory+data+analysis>
- Use the above link to search for exploratory data analysis notebooks.
- Practice EDA using at least 3 datasets.
  - e.g. <https://www.kaggle.com/datasets/rishabhkarn/ipl-auction-2023/data>

- **Assignment**

- ☐ Perform EDA (Exploratory data analysis on **at least 2 additional datasets** on Kaggle)

## Week 18, 19, 20, 21: Machine Learning 🛠️

---

- **Machine Learning: Preprocessing**

- Handling NA values, outlier treatment, data normalization
- One hot encoding, label encoding
- Feature engineering
- Train test split
- Cross validation

- **Machine Learning: Model Building**

- Types of ML: Supervised, Unsupervised
- Supervised: Regression vs Classification
- Linear models
  - Linear regression, logistic regression
  - Gradient descent
- Nonlinear models (tree-based models)
  - Decision tree
  - Random forest
  - XGBoost
- Model evaluation
  - Regression: Mean Squared Error, Mean Absolute Error, MAPE
  - Classification: Accuracy, Precision-Recall, F1 Score, ROC Curve, Confusion matrix
- Hyperparameter tuning: GridSearchCV, RandomSearchCV



- Unsupervised: K means, Hierarchical clustering, Dimensionality reduction (PCA)
- **Learning Resources**
  - Track A
    - YouTube playlist (more than 2 million views): <https://bit.ly/3io5qqX>
    - First 21 videos
    - Feature engineering playlist: <https://bit.ly/3lFa3Yf>
  - Track B (Affordable Fees)
    - Master Machine Learning for Data Science & AI: This course takes you from beginner to advanced levels, providing deep intuition on algorithms, engaging cinematic experiences, end-to-end projects, and hands-on coding practice: <https://codebasics.io/courses/machine-learning-for-data-science-beginners-to-advanced>
- **Core/Soft Skills**
  - **Project Management**
    - Scrum: <https://scrumtrainingseries.com/>
    - Kanban: <https://youtu.be/jf0tlbt9lx0>
    - Tools: JIRA, Notion
- **Assignment**
  - ☐ Complete all exercises in ML playlist: <https://bit.ly/3io5qqX>
  - ☐ Work on **2 Kaggle ML notebooks**
  - ☐ Write **2 LinkedIn posts** on whatever you have learnt in ML
  - ☐ Discord: Help people with **at least 10 answers**
  - ☐ Track B: Finish exercises and quizzes for relevant topics

## Week 22: ML Ops

---

- **Topics**
  - What is API? FastAPI for Python server development
  - DevOps Fundamentals: CI/CD pipelines, containerization (Docker, Kubernetes)
  - Familiarity with at least one cloud platform (AWS, Azure etc.)
- Learning Resources
  - Track A:
    - FastAPI tutorial: <https://bit.ly/497p6Ex>
    - Docker tutorial: <https://bit.ly/3uCNpeE>

- Track B (Affordable Fees):
  - Included in the above Master Machine Learning for Data Science & AI

## Week 23, 24: Machine Learning Projects with Deployment

---

- You need to finish **two** end to end ML projects. One on **Regression**, the other on **Classification**
- Regression Project: Bangalore property price prediction
  - YouTube playlist link: <https://bit.ly/3ivycWr>
  - Project covers following
    - Data cleaning
    - Feature engineering
    - Model building and hyper parameter tuning
    - Write flask server as a web backend
    - Building website for price prediction
    - Deployment to AWS
- Classification Project: Sports celebrity image classification
  - YouTube playlist link: <https://bit.ly/3ioaMSU>
  - Project covers following
    - Data collection and data cleaning
    - Feature engineering and model training
    - Flask server as a web backend
    - Building website and deployment
- **ATS Resume Preparation**
  - Resumes are dying but not dead yet. Focus more on online presence.
  - Here is the resume tips video along with some templates you can use for your data analyst resume: <https://www.youtube.com/watch?v=buQSI8NLOMw>
  - Use this checklist to ensure you have the right ATS Resume: [Check here.](#)
- **Portfolio Building Resources:**

You need a portfolio website in 2024. You can build your portfolio by using these free resources.

  - [GitHub](#)
    - Upload your projects with code on github and using github.io create a portfolio website
    - Sample portfolio website: <http://rajag0pal.github.io/>

- [Linktree](#)
  - Helpful to add multiple links in one page.
- **Assignment**
  - In above two projects make following changes
    - ☐ Use **FastAPI** instead of **flask**. FastAPI tutorial: <https://youtu.be/Wr1JjhTt1Xg>
    - ☐ **Regression project**: Instead of property prediction, take any other project of your interest from Kaggle for regression
    - ☐ **Classification project**: Instead of sports celebrity classification, take any other project of your interest from Kaggle for classification and build end to end solution along with **deployment to AWS or Azure**
    - ☐ Add a link of your projects in your resume and LinkedIn.

(Tag Codebasics, Dhaval Patel and Hemanand Vadivel with the hashtag #dsroadmap24 so we can engage to increase your visibility)

## Week 25, 26, 27: Deep Learning 🤖

---

- **Topics**
  - What is a neural network? Forward propagation, back propagation
  - Building multilayer perceptron
  - Special neural network architectures
    - Convolutional neural network (CNN)
    - Sequence models: RNN, LSTM
- **Learning Resources**
  - Deep Learning playlist (tensorflow): <https://bit.ly/3vOZ3zV>
  - Deep learning playlist (pytorch): <https://bit.ly/3TzDbWp>
  - End to end potato disease classification project: <https://bit.ly/3QzkVJi>
- **Assignment**
  - ☐ Instead of potato plant images use tomato plant images or some other image classification dataset.
  - ☐ Deploy to Azure instead of GCP.
  - ☐ Create a presentation as if you are presenting to stakeholders and upload video presentation on LinkedIn.

## Week 28, 29, 30: NLP or Computer Vision & GenAI

---

- Many AI engineers choose a specialized track which is either NLP or Computer vision. You don't need to learn both.
- **Natural Language Processing (NLP)**
  - Topics
    - Regex
    - Text presentation: Count vectorizer, TF-IDF, BOW, Word2Vec, Embeddings
    - Text classification: Naïve Bayes
    - Fundamentals of Spacy & NLTP library
    - One end to end project
  - Learning Resources
    - NLP YouTube playlist: <https://bit.ly/3XnjfEZ>
- **Computer Vision (CV)**
  - Topics
    - Basic image processing techniques: Filtering, Edge Detection, Image Scaling, Rotation
    - Library to use: OpenCV
    - Convolutional Neural Networks (CNN) – Already covered in deep learning.
    - Data preprocessing, augmentation – Already covered in deep learning.
- **Assignment**
  - NLP Track: Complete exercises in this playlist: <https://bit.ly/3XnjfEZ>

## Week 31, 32: LLM & Langchain

---

- Topics
  - What is LLM, Vector database, Embeddings?
  - RAG (Retrieval Augmented Generation)
  - Langchain framework
- Learning Resources
  - Langchain, LLM playlist: <https://bit.ly/3RYpxuw>

## Week 33 onwards.... 😊😊😊

---

- More projects 🧠
- Online brand building through LinkedIn, Kaggle, Discord, Opensource contribution 👥
- Job application and Success 🚀

## Tips of effective learning 🔥

---

- **Spend less time in consuming information, more time in**
  - Digesting
  - Implementing
  - Sharing
- **Group learning**
  - Use **partner-and-group-finder** channel on codebasics discord server for group study and hold each other accountable for the progress of your study plan. Here is the discord server link: <https://discord.gg/r42Kbuk>