



# AI Champ Help Book

An Initiative by Codevector Labs

Email: **[aichamp@codevector.in](mailto:aichamp@codevector.in)**

Contact: **8777239040 9407608477**

Website: **<https://codevector.in>**

Kolkata 700054



# TABLE OF CONTENTS

Introduction	3
Prerequisites	4
Python (Required)	4
ML Basics (Required)	4
Linux (Optional)	4
If the candidate is new	4
Course Expectations	5
Directly	5
Indirectly	5
What Not To Expect?	6
Syllabus	7
Deep Learning Basics	7
Computer Vision Basics	7
Natural Language Processing Basics	8
Case Studies	8
Selection Procedure	9
Champ Marathon (Round 1)	9
Scoring for Screening Test	9
Final Interview (Round 2)	10
Why not MCQs?	10
Stay Updated	11
Follow Us on Social Media	11
POC	11
Appendix A	12
Resources for Python	12
Resources for ML (Basics)	12
Resources for Linux	12

# INTRODUCTION



The course aims to groom the students as a good programmer who can **learn new things** quickly and add new skills to the arsenal. The course uses only the **content that is freely available** on the internet, with some exceptions of (optional) book recommendations. Just writing the right code is not the end of the world. For this reason, this course also consists of some modules which are helpful to **showcase work** by writing right notes, blogs, or posts on LinkedIn, GitHub. In terms of AI, the course consists of many different real-life case studies that the students need to solve. The problems aim to make students better at **Deep Learning, Computer Vision, and Natural Language Processing**. A **discord community** consisting of students and the **Guides/TAs** are available to solve the problems that one might face throughout the course.



***“Don’t be pushed around by the fears in your mind. Be led by the dreams in your heart.”***

— Roy T. Bennett, *The Light in the Heart*

# PREREQUISITES

## Python (Required)

We suggest the students have some hands-on with Python. Python is a straightforward language to learn yet powerful.

## ML Basics (Required)

One should know the basics of Machine Learning. Basics are required. For example, the candidate must be able to answer what is a loss function, gradient descent, confusion matrix. Well, if these terms are new to the reader, we recommend them to read about it. There are plenty of free resources available online.

## Linux (Optional)

The participant should know the basic Linux commands. It is always good to use a Linux based operating system for programmers. However, it is entirely optional, and the candidate can go with Windows/macOS if they are very sure that they may solve the problems that might occur due to different Operating Systems. For example, many tutorials and code repositories assume that they are using Linux and have little support for other Operating Systems.

## If the candidate is new

If the candidate just started these things and is worried whether he/she will be able to do it or not, then let us assure that if anyone may do it, they too may. They just need a willingness to do it. It is also never too late to start. We have attached all the resources which will help them in quick revision or quick start. Please scroll to the last of this document, and in Appendix A, we have attached all such resources for each of the topics above.

# COURSE EXPECTATIONS

Note: **The course is not limited to a set of things to be learned. The whole idea of this course is to enhance the problem-solving skills of the student by providing practical tips and methods.**

## Directly

1. Real-world case studies where the problems can be solved using some domain of AI.
2. A discord community of the students, along with the Guides/TAs.
3. Open source content (blogs, videos, articles), curated to help them learn and understand the basics of Computer Vision, Deep Learning, and Natural Language Processing.
4. A document (unique for each case study) which will help you understand how to solve the real-world case studies technically, without spoilers!
5. Practical hands-on mini-projects and real-world case studies.
6. Guidance on showcasing their projects on LinkedIn, GitHub, Blogs.
7. A Guide/TA to explain complex information and clear their doubts.
8. Practical datasets, which are a mix of publicly available datasets and custom, created ones for the course.

## Indirectly

1. A better GitHub Profile.
2. Good theoretical as well as practical knowledge of CV, NLP, and DL.
3. Write blogs that are effective and helpful for the community
4. Peer-based learning.
5. Learn to read code, error messages, and fix them.
6. Learn to read the documentation of many different (generally open-source) software and use them.
7. Find out the challenges in real-world projects and try to fix them as much as possible.
8. Practical tips to approach and solve the problems.

**Learning by Doing** is the main essence of this course; first, the participants try to read those resources. If they have any doubts or confusion, they may ping any mentor specified. The mentor will help the candidate with those topics.

**Hands-On** experiences are top on the jerry. We have curated a curated case-studies from the projects that we have made and deployed in real life in the past few years.

# WHAT NOT TO EXPECT?

Note: **This is not a traditional course. There are neither pre-recorded videos nor live coding sessions. Instead of teachers, we have guides who will help the participant solve their problems and explain complicated things when required.**

We will be having a discord server for the community forum and announcements. Students may put their queries directly on the discord, and the mentor will help the candidate with any of the issues. One question may arise, why there are no teacher teaching concepts; the reason is that it is an old way of teaching. Now, in the digital era, we believe that there is plenty of information out there. All the candidate needs is a guide who may guide, how to look for that information on the internet, and how not to.

All the mentors are experts in this field with some hands-on industry experience. Moreover, we have made sure that they have the required teaching skills, as well. In the whole course duration, we will work on topics specified in the Syllabus section, along with hands-on some of the curated industry Use-Cases that we have prepared.

This course is different in terms of many aspects. First of all, this will not be like a lecture where a tutor will be teaching some concepts. Instead, we believe that everyone has a different pace of learning. The student will be provided with a source either created by the Codevector Labs or any 3<sup>rd</sup> party such as blogs, videos, gists.

# SYLLABUS

**Note: The course syllabus depends on what real-world case study the student chooses to solve. By the end of the course, the student will understand the basics of Computer Vision, Deep Learning, and Natural Language Processing. The following information about the syllabus is just to give students an idea of what they are going to learn but not limited to it. The approach is more practical and not too theoretical.**

## Deep Learning Basics

1. A brief history of Deep Learning and inspiration from the human brain.
2. Overview of mathematics which will help students understand the language used in deep learning.
3. ANN and its components - Perceptron, Layers, Weight, Bias, Loss, Backpropagation, Activation Functions.
4. Mini project - Training a deep learning model with Keras/Tensorflow/Pytorch on a standard dataset.
5. Understanding the general problems and some tips to train the models – Imbalanced Datasets, Overfitting, Underfitting, Regularization, Augmentation, Small Datasets, Transfer Learning, Vanishing Gradients, Exploding Gradients, Normalization.

## Computer Vision Basics

1. A brief history of Computer Vision and inspiration from the human vision.
2. Overview of mathematics will help you understand the language used in computer vision. It mostly contains matrix operations.
3. Understanding what an image is to a computer, the coordinate system of the image, overview of some operations performed on the image/videos.
4. Mini project - Image processing exercise with NumPy and OpenCV. Understand popular kernels - Sobel, Laplacian, Gaussian.
5. Mini project - Background Subtraction, Motion Detection, and Contour Detection.
6. CNN and its components - Similarity with ANN, Convolution operation, Kernels, Activation Maps, Pooling, Padding, Flattening, FCN.
7. Mini project - Training a deep learning model (CNN) with Keras/Tensorflow/Pytorch on a standard dataset.
8. Understanding the general problems and some tips to train the model - Same as deep learning but more specific to CNNs. Eg. Different Augmentation techniques, Using different architectures of CNN models for their use case, BatchNormalization.

## Natural Language Processing Basics

1. Can machines understand the natural language as humans do?
2. Do computers only understand numbers? How can computers represent natural language in numbers?
3. NLP applications, overview, and pipeline - Sentence Segmentation, Word Tokenization, POS Tagging, Lemmatization, Stemming, Stop Words, NER.
4. Mini Project - Building a web scraper using python.
5. Using popular NLP libraries - NLTK, SpaCy, Gensim, Transformers.
6. Mini Project - Building a simple document summarizer
7. Overview of deep learning methods for NLP tasks.

## Case Studies

1. 10+ real-world case studies - select one for the main capstone project.
2. Case-study specific curated learning content - Blogs, Videos, Slides, Research Papers.
3. Tips on how to find current state-of-the-art (SOTA) methods for a given task.
4. Tips on using open-source libraries and codes for the project.
5. Datasets - Publicly available, exclusively created datasets.
6. Tips on creating custom datasets - How to reduce efforts while creating a dataset?
7. Overview of some software and platforms used to annotate the datasets.
8. A model works decently on the test dataset but does not work very well on real-world,
  1. uncontrolled data. What to do now?
9. What are REST APIs?
10. Creating WEB APIs using python frameworks such as Flask, Django, FastAPI.
11. Deploy the AI/ML models to the cloud.
12. Understanding the cost required to deploy the AI/ML models.
13. Perform load testing on the deployed solution - How much load can it handle?
14. Can we optimize to reduce latency and cost?
15. How will we know if our server has crashed? Do we have any backup plans? How to fix it ASAP?
16. Our two cents on - What to do next? How to become an expert in the domain of their choice?



# SELECTION PROCEDURE

## Champ Marathon (Round 1)

We will conduct a five-day selection process in the format of a community based technical marathon **from 24th August 2020 to 28th August 2020**.

- On **24th August**, we shall **release a set of four simple tasks**.
- We will add all the applicants **to a community before 24th august**, where they can start conversations with the other candidates.
- These four tasks will range from **activities like dataset collection, straightforward coding to finally building a small application**.
- A candidate doesn't need to attempt **all tasks**. Each of the four tasks will have a specific score and the total aggregate score shall be 100
- A candidate can **discuss in the community, refer to the internet or any other sources for the solution**; however, **submission has to be at an individual level**.
- The candidate shall **receive hints from the Codevector Labs team** throughout the four days on how to solve each activity; we will try to make the entire marathon a fun and learning experience.
- Candidates have to **submit their complete solution** for all four tasks by **29th August 11:59 PM**.
- The **minimum total score that a candidate needs to score is 50 (out of 100)** for the interview. Therefore, a candidate can attempt two out of four tasks and still be eligible for interview round.
- Based on the **score secured by candidates in the technical marathon and the online interview**, we shall select the most deserving candidates to be a part of AI champ Batch 1.
- We will reject both the candidates; in case we find **plagiarism between the submission of two candidates** for further consideration (we will not reach out for any clarification from either candidate).
- It is a **community-based selection process**, and we shall choose the deserving candidates for the next round, which is an online interview.

## Scoring for Screening Test

If you don't have a **GitHub username**, then it's high time that you should make one! Each of the four tasks shall have a specific score, and the **aggregate score of all the four tasks shall be 100**.

The most critical parameters while scoring is:

1. Frequency of commits on GitHub
2. Quality of your code
3. The way you demonstrate your solution

#### 4. How well you maintain your documentation

The scoring on each task shall be done **unbiased** without considering the candidate's educational or professional background. We will evaluate based on What you submit at the end!

### Final Interview (Round 2)

We usually take group interviews, where each candidate can hear other answers, and every candidate will get an equal opportunity to speak. The primary objective of the meeting is to check the curiosity of an applicant to learn AI.

Don't worry much about the details about the interview and the tasks. If you feel that you probably won't make it, we want to tell you that you will likely make it through if you give it a try. The interview process will also be straightforward and fun!

If you are excited about the first of its kind technical marathon, let us inform you that the **AI champ team is here to help you!** We will try our best to support you throughout the process and help you learn something exciting!

We usually take group interviews, where each candidate can hear other answers, and every candidate will get an equal opportunity to speak. The primary objective of the meeting is to check the curiosity of an applicant to learn AI.

Don't worry much about the details about the interview and the tasks. If you feel that you probably won't make it, we want to tell you that you will likely make it through if you give it a try. The interview process will also be straightforward and fun!

If you are excited about the first of its kind technical marathon, let us inform you that the AI **champ team is here to help you!** We will try our best to support you throughout the process and help you learn something exciting!

### Why not MCQs?

We do not think some set of predefined questions may answer our requirements. Considering there is an excellent probability that the answer is available in some Google Search. All the candidates are unique and have a unique set of skills that otherwise may not be judged based on some technical MCQs.

# STAY UPDATED

## Follow Us on Social Media

**Facebook:** <https://www.facebook.com/TeamCodeVector/>

**Twitter:** <https://twitter.com/codevectorlabs>

**YouTube:** <https://www.youtube.com/channel/UCM-Gwf-LXux5LhwciWAXw-Q>

**Instagram:** <https://www.instagram.com/codevectorlabs/>

**LinkedIn:** <https://www.linkedin.com/company/codevector-ai-labs/>

**Email:** [aichamp@codevector.in](mailto:aichamp@codevector.in)

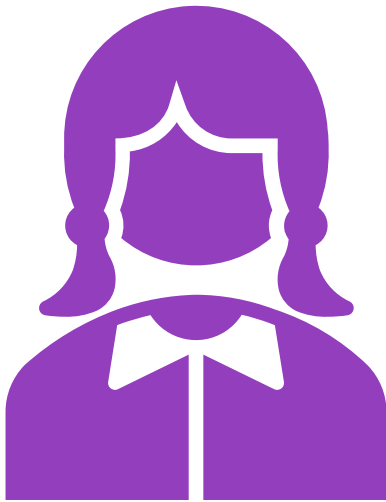
## Join Discord

All our communications will be done via Discord, we request you to join the discord server.

**Join Server:** <https://discord.gg/zEWddDH>

## POC

Reach out to us for any queries.



**Aliyasmin Rahaman**

+91 87772 39040 | [LinkedIn](#)



**Kuldeep Pisda**

+91 94076 08477 | [LinkedIn](#)

# APPENDIX A

## Resources for Python

1. <https://www.geeksforgeeks.org/introduction-to-python-for-absolute-beginners/>
2. <https://www.python.org/about/gettingstarted/>
3. <https://www.learnpython.org/>
4. <https://www.w3schools.com/python/>

## Resources for ML (Basics)

1. <https://towardsdatascience.com/introduction-to-machine-learning-for-beginners-eed6024fdb08>
2. <https://www.analyticsvidhya.com/blog/2019/08/detailed-guide-7-loss-functions-machine-learning-python-code/>
3. <https://www.geeksforgeeks.org/ml-linear-algebra-operations/>
4. <https://towardsdatascience.com/gradient-descent-explained-9b953fc0d2c>
5. <https://machinelearningmastery.com/gradient-descent-for-machine-learning/>
6. [https://www.youtube.com/watch?v=jGwO\\_UgTS7I&list=PLoROMvodv4rMiGQp3WXShtMGgzqpfVfbU](https://www.youtube.com/watch?v=jGwO_UgTS7I&list=PLoROMvodv4rMiGQp3WXShtMGgzqpfVfbU)

## Resources for Linux

1. <https://www.pcsuggest.com/basic-linux-commands/>