Course Syllabus: Vision AI Developer

Course Title: Seeing the World Through AI: Mastering Computer Vision

Target Audience: Suitable for developers, engineers, data scientists, and students interested in AI for image and video analysis. Basic programming knowledge (e.g., Python) and interest in visual data are helpful but not required.

Course Level: Comprehensive program covering Basic, Intermediate, and Advanced levels. Duration: 12 weeks (flexible for self-paced learning).

Course Description:
This course trains students to become Vision AI Developers, specializing in AI for computer vision tasks like image recognition and object detection. You'll learn to build systems that analyze images and videos, such as identifying menu items for a platform like Zomato. From foundational image processing to advanced deep learning models, you'll develop skills to create vision-based AI solutions that enhance user experiences and business operations.

Learning Objectives:

Upon completion, students will be able to:

- Understand computer vision concepts and techniques.

- Process and analyze images/videos using Python and AI frameworks.

- Build and train models for image recognition, object detection, and segmentation.

- Integrate vision AI into real-world applications (e.g., food delivery apps).

- Address ethical and privacy issues in vision AI.

- Develop a portfolio of vision AI projects.

Course Structure:

Part 1: Basic Foundations (Weeks 1-4)

This section introduces computer vision and image processing basics.

- Week 1: Introduction to Vision AI

    - What is computer vision? Image recognition, object detection, segmentation.

    - Role of a Vision AI Developer.

    - Examples: Identifying dishes in Zomato's menu photos.

    - Exercise: Explore a vision AI tool (e.g., Google Vision API).

- Week 2: Image Processing Basics

    - Image fundamentals: Pixels, color spaces, resolution.

    - Python tools: OpenCV, PIL for image manipulation.

    - Hands-on: Apply filters to a sample image (e.g., restaurant dish).

- Week 3: Machine Learning for Vision

    - Supervised learning for images: Classification, regression.

    - Datasets: ImageNet, COCO for training models.

    - Exercise: Train a simple classifier to recognize food images.

- Week 4: Introduction to Deep Learning

    - Neural networks for vision: Convolutional Neural Networks (CNNs).

    - Frameworks: TensorFlow, PyTorch.

    - Hands-on Project: Build a model to classify restaurant photos (e.g., pizza vs. burger).

Part 2: Intermediate Concepts (Weeks 5-8)

This section focuses on advanced vision tasks and model training.

- Week 5: Object Detection

    - Techniques: YOLO, Faster R-CNN for detecting objects.

    - Use case: Detecting items in a restaurant order image.

    - Hands-on: Implement an object detection model.

- Week 6: Image Segmentation

  - Semantic and instance segmentation: Mask R-CNN.

  - Applications: Isolating food items in photos for Zomato's menu.

  - Exercise: Segment objects in a sample food image.

- Week 7: Model Evaluation and Optimization

  - Metrics: mAP, IoU for vision tasks.

  - Optimizing models: Transfer learning, data augmentation.

  - Hands-on: Fine-tune a pre-trained model for a specific task.

- Week 8: Real-World Integration

  - Deploying vision models: APIs, cloud services (e.g., AWS Rekognition).

  - Case Study: Vision AI for Zomato's food quality checks.

  - Hands-on Project: Build and deploy a vision AI app for food classification.

Part 3: Advanced & Expert-Level Application (Weeks 9-12)
This section prepares students for enterprise-grade vision AI solutions.

- Week 9: Advanced Vision Models

  - Transformers for vision: Vision Transformers (ViT), Swin Transformers.

  - Multimodal models: Combining images and text (e.g., CLIP).

  - Exercise: Build a model combining text and images for menu analysis.

- Week 10: Real-Time Vision AI

  - Real-time processing: Video analysis, edge deployment.

  - Tools: OpenVINO, TensorRT for optimization.

  - Hands-on: Deploy a real-time object detection system.

- Week 11: Ethics and Privacy in Vision AI

  - Ethical issues: Bias in facial recognition, privacy in image data.

  - Compliance: GDPR for visual data.

  - Exercise: Audit a vision AI system for ethical risks.

- Week 12: Capstone Project & Trends

    - Capstone Project: Develop a vision AI system for a Zomato-like platform (e.g., dish recognition for menu uploads).

    - Trends: AR/VR vision AI, 3D object detection.

    - Career paths: Vision AI development, autonomous systems, tech startups.

Assignments & Grading:

- Weekly Coding & Image Labs: 25%

- Intermediate Projects (Weeks 4 & 8): 30%

- Capstone Project: 35%

- Class Participation & Peer Reviews: 10%