Course Type	Course Code	Name of Course	L	T	P	Credit
DE	OCSD604	Foundations of Deep Learning: Concepts and Applications	3	0	0	3

Course Objective
Deep Learning is a core pillar of modern Artificial Intelligence, powering applications in computer vision, natural language processing, healthcare, and robotics. With the rapid expansion of AI-related programs in AICTE-affiliated institutions, it's essential for students to build a strong foundation in this field.
Learning Outcomes
This course is designed to guide learners' step-by-step—from the basics of neural networks to advanced architectures like CNNs, RNNs, and Autoencoders. Emphasis is placed on both conceptual clarity and practical implementation using Python and Google Colab. By the end of the course, students will be able to apply deep learning algorithms on real-world data to develop intelligent solutions.

Unit No.	Topics To be covered	No of Lecture Video	Video Time (Minutes)	Learning Outcome
1	Overview and motivation for the course Why DL, Importance, Companies which are working, Applications, Future. To whom the course is designed (target audience), Contents (week wise). How this course is different from others. (More hands-on) and live classes, Make it more impactful. Overview of machine learning and deep learning, difference between ML and DL with an example, History and Evolution of Deep Learning with computational efficiency. Introduction to Neural networks: Perceptron: logistic regression, Single Layer Perceptron, Single Layer Perceptron numerical problem, Limitations of Single Layer Perceptron Hands-on on building a simple perceptron model using colab file. Introduction to Multilayer Perceptron, Difference between Shallow neural networks, deep neural networks, take an example and show how to design the NN (5 to 7 examples), Activation Functions, Loss Functions	5	206	Understand the strategic importance, industrial applications, and historical evolution of Deep Learning, distinguishing its capabilities and computational efficiency from traditional Machine Learning. Master the fundamentals of Single-Layer Perceptrons and logistic regression, solving numerical problems while analyzing structural limitations like linear separability. Build and implement a functional Single-Layer Perceptron model from scratch through hands-on live coding sessions using Google Colab. Differentiate between shallow and deep neural network architectures while designing Multilayer Perceptron (MLP) structures for diverse real-world use cases. Select and apply appropriate activation functions and loss functions to optimize gradient flow, model training, and network performance.
2	Gradient Descent (GD) and Backpropagation (MSE) Optimizers: Momentum-Based GD, Nesterov Accelerated GD, Stochastic GD, AdaDelta, AdaGrad, RMSProp, Adam. Regularization Techniques: L1/L2 regularization, dropout, Early stopping. Hands-on on building the Artificial Neural Network for classification and regression problems with exposure to hyperparameter tuning. Interpreting the results using simple XAI techniques: LIME & SHAP.	5	234	Master the mechanics of Gradient Descent and Backpropagation using Mean Squared Error (MSE) to compute loss gradients and update neural network weights efficiently. Analyze and apply diverse optimization algorithms—including SGD, Momentum, Nesterov, AdaGrad, AdaDelta, RMSProp, and Adam—to improve convergence speed and model stability. Implement L1/L2 regularization, dropout, and early stopping to prevent overfitting and improve neural network generalization. Construct hands-on Artificial Neural Network (ANN) architectures for both classification and regression tasks while conducting systematic hyperparameter tuning. Interpret network predictions and feature importance using Explainable AI (XAI) frameworks like LIME and SHAP to ensure model transparency and trustworthiness.
3	CNN: Fundamentals of Image representation and Image preprocessing and Data augmentation Introduction to Convolutional Neural Networks, Inspiration behind CNN, Key Components of CNN, Types of convolutions CNN architecture	5	300	Master digital image representation, preprocessing pipelines, and data augmentation techniques to improve model generalization and mitigate overfitting. Explain the foundational concepts and biological inspiration behind Convolutional Neural Networks (CNNs), identifying core components like

	Hands-on on building a simple CNN model for binary and multiclass classification.			convolutional, pooling, and fully connected layers. Analyze diverse convolution operations and architectural design principles used in modern CNNs for hierarchical spatial feature extraction. Construct and train a custom CNN model from scratch for binary image classification problems using hands-on deep learning workflows. Implement and evaluate end-to-end CNN architectures for multiclass classification tasks, optimizing loss functions, hyperparameters, and performance metrics.
4	A typical CNN structure, Standard CNN models: AlexNet, VGGNet 16 and 19 Standard CNN Models: GoogLeNet, ResNet 18, 34 Standard CNN Models: Inception, Transfer Learning Hands-on on transfer learning and building an ensemble mode	6	189	Analyze standard Convolutional Neural Network architectures—including AlexNet, VGGNet (16/19), GoogLeNet, ResNet (18/34), and Inception—evaluating structural innovations like residual shortcuts and multi-scale filters. Apply transfer learning strategies by adapting pre-trained vision models for custom feature extraction, fine-tuning, and domain adaptation tasks. Construct and evaluate ensemble models combining multiple fine-tuned CNN architectures to maximize classification accuracy and generalizability through hands-on implementation.
5	Introduction to XAI: Algorithms and its working mechanism Hands-on: Interpreting the results from CNN model using simple XAI techniques: GRADCAM and SMOOTHGRAD	5	230	Understand foundational Explainable AI (XAI) concepts, algorithms, and working mechanisms to interpret complex deep learning model decisions. Analyze the mechanics of gradient-based visual explanation techniques, focusing on class-activation mapping via Grad-CAM and noise-reduction through SmoothGrad. Implement hands-on workflows to extract, visualize, and interpret visual saliency maps from Convolutional Neural Networks (CNNs) using Grad-CAM and SmoothGrad.
6	Evaluation metrics for segmentation, CNN based segmentation algorithms UNet: Attention-based UNet, Introduction to CNN based Object Detection models. Object detection algorithms: YOLO, RCNN, Faster RCNN models. Hands-on on object detection using YOLO. Hands-on on UNet and attention based UNet	7	235	Evaluate semantic image segmentation using key evaluation metrics (IoU, Dice coefficient) while analyzing CNN-based architectures including UNet and Attention-based UNet. Analyze the mechanics of foundational object detection frameworks, contrasting two-stage region proposal networks (R-CNN, Faster R-CNN) with single-stage real-time detectors (YOLO). Build, train, and evaluate end-to-end computer vision pipelines through hands-on practice using YOLO for object detection and UNet/Attention-UNet for image segmentation.
7	Sequence-to-sequence models Introduction to Recurrent Neural Networks and their structure Challenges in RNN (Vanishing and Exploding Gradients). Numerical problem on RNN. Hands-on on building RNN on structured and unstructured data. Variants of RNN and its hands-on	5	272	Analyze sequence-to-sequence modeling and Recurrent Neural Network (RNN) architectures, addressing structural limitations like vanishing and exploding gradients. Solve step-by-step numerical problems to calculate hidden state transitions, activation outputs, and gradient propagation in RNN cells. Build, train, and evaluate standard RNNs alongside advanced variants (such as LSTM and GRU) on structured and unstructured sequential data through hands-on implementation.
8	Introduction to Long-short term memory (LSTM) architecture and its necessity, Bidirectional LSTM, Stacked LSTMs Understand the GRU architecture, Compare LSTM vs GRU: speed, accuracy,	5	164	Analyze advanced LSTM architectures (including Bidirectional and Stacked variants) alongside GRU trade-offs—evaluating speed, accuracy, and computational complexity to select the optimal model.

	complexity. When to use GRU over LSTM. Why attention mechanism in RNNs, working of attention mechanism, benefits of attention mechanism Hands-on on building LSTM models for structured and unstructured data.			Explain the necessity, mechanics, and benefits of the attention mechanism in sequence models for capturing long-range dependencies and resolving context bottlenecks. Construct, train, and evaluate LSTM models on both structured time-series and unstructured sequential data through hands-on deep learning workflows.
9	Introduce NLP tasks, Classical NLP vs Deep Learning NLP, Text Preprocessing Tokenization, Stopwords, Lemmatization. Word Representations: One-hot encoding, Word embeddings: Word2Vec, GloVe, FastText. Sequence modeling in NLP, Recurrent Neural Networks (RNN) basics, Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU), Word embeddings + RNN for sequence tasks. Hands on session on RNN for NLP task	5	165	Compare classical and deep learning NLP paradigms while implementing core text preprocessing pipelines, including tokenization, stopword removal, and lemmatization. Evaluate word representation techniques, contrasting sparse one-hot encoding with dense vector embeddings such as Word2Vec, GloVe, and FastText for capturing semantic relationships. Construct and train sequence models combining word embeddings with RNN, LSTM, and GRU architectures to solve NLP tasks through hands-on implementation.
10	Unsupervised Learning Introduction to Autoencoder, Architecture of AE and math behind. Types (Simple, Deep, CNN-based, Types), Training of Autoencoders, Hands-on on building a AE and types of AE.	6	250	Analyze the core architecture, mathematical formulation, and reconstruction loss mechanics behind Autoencoders for unsupervised representation learning. Differentiate between diverse Autoencoder variants—including Simple, Deep, and Convolutional Autoencoders (CAEs)—for feature extraction and spatial representation. Construct and train custom Autoencoder architectures from scratch through hands-on implementation to solve unsupervised representation and data reconstruction tasks.
11	Transformer architectures Self-Attention, Encoder Decoder Attention, In-context-learning, Low-rank adaptation. Self-supervised learning: Objectives and Loss Functions, Masked Language Modeling.	6	198	Master the core mechanics of Transformer architectures, including Self-Attention, Multi-Head Attention, and Encoder-Decoder cross-attention mechanisms. Analyze self-supervised pre-training frameworks, loss functions, and objectives like Masked Language Modeling (MLM) alongside fine-tuning paradigms like In-Context Learning and Low-Rank Adaptation (LoRA). Implement and evaluate Transformer-based models to solve downstream natural language processing tasks efficiently using parameter-efficient fine-tuning strategies.
12	Large Language Models Tokenizers, Pre-training and post-training, multimodal alignment, model compression, Reinforcement Learning for fine-tuning, Proximal Policy Optimization, Benchmarking and Evaluation of LLMs. Diffusion Models: Deep generative models, VAEs and GANs, Forward and reverse diffusion, Denoising Score Matching, Variational Lower Bounds, Stable Diffusion	5	174	Analyze the complete lifecycle of Large Language Models—including subword tokenization, pre-training, post-training alignment (RLHF/PPO), multimodal grounding, model compression, and standardized evaluation benchmarks. Evaluate deep generative modeling paradigms by contrasting Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs) with forward and reverse Diffusion processes. Formulate the mathematical and operational mechanics of latent diffusion models, focusing on Denoising Score Matching, Variational Lower Bounds (ELBO), and cross-attention architectures in Stable Diffusion.
Total No of video lectures & Video time		**65**	**2617**	

Text books

1.Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. Deep learning. MIT Press, 2016, ISBN : 9780262035613.

2.Zhang, Aston, et al. "Dive into deep learning." Cambridge University Press, 2023, ISBN-13 : 9781009389433.

Reference books:

1.CS231n: Deep Learning for Computer Vision, Stanford University.

2.Practical Deep Learning, Fast.ai (https://course.fast.ai/)

Link of NPTEL Course:

https://nptel.ac.in/courses/106108840