AI Research Internship and Training Program

About

This internship is designed to provide students with practical exposure to Artificial Intelligence and Machine Learning through a structured, hands-on learning approach. Participants will start by building a strong foundation in data analysis and machine learning, gradually progressing to deep learning, computer vision, natural language processing, and modern AI systems such as transformers and Retrieval Augmented Generation (RAG). By the end of the program, students will have developed multiple working projects and a solid understanding of how AI systems are designed and applied in real-world applications and authorized certificate of internship from Thundertribes.

Contents

Module 1 - "When Data Starts Talking"

- Environment setup (Python, Jupyter, libraries)
- Intro to NumPy, Pandas, Matplotlib
- Data cleaning & preprocessing
- Regression & Classification basics
- Train Linear/Logistic + Random Forest models
- Model evaluation

Module 2 - "How Machines Learn to Think"

- Cross-validation & overfitting
- K-Means clustering
- PCA for visualization
- Error analysis
- Feature engineering

Hands On Project:

Mini ML pipeline project

Module 3 - Deep Learning & PyTorch Core

- Neural network basics (neurons, activations, layers)
- Intro to tensors & PyTorch
- Autograd & backpropagation (conceptual)
- nn.Module structure

- Training & validation loops
- Build a simple NN on MNIST

Outcome:
Students can build and train a neural network from scratch.

Module 4 - When Machines Learn to See

- Overfitting & regularization (Dropout, BatchNorm)
- Learning rate & early stopping
- CNN intuition (convolutions, feature maps, pooling)
- Build CNN for image classification
- Transfer learning (ResNet pretrained)
- Compare scratch vs pretrained model
- Basic error analysis

Hands-On Project:
Build an image classification system and improve its performance.

Module 5 - "Teaching Machines to Understand Language"

Students are introduced to Natural Language Processing and how machines understand text.

Topics Covered:

- What is NLP?
- Text preprocessing (Tokenization, Stopwords, Stemming, Lemmatization)
- Bag of Words (BoW)
- TF-IDF & N-grams
- Feature vectors
- Limitations of TF-IDF
- Word Embeddings (Word2Vec, GloVe)

- Neural Network basics
    - Perceptron concept
    - Activation functions (ReLU, Sigmoid)
    - Loss function intuition
    - High-level training process

Hands-On:
Build a basic text classification system using TF-IDF and a simple neural network.

Module 6 - "History Lessons - every fan should know"

Students explore how modern NLP models capture context and meaning.

Topics Covered:

- RNN & LSTM (conceptual understanding)
- Why RNN fails for long sequences
- Attention mechanism
- What is a Transformer?
- Encoder vs Decoder architecture
- BERT vs GPT
- Fine-tuning pretrained models

Hands-On:
Fine-tune a pretrained Transformer (BERT) for text classification and compare with classical NLP approaches.

Module 7 - "Building an AI That Knows Your Data"

Students learn how modern AI systems retrieve and generate knowledge instead of relying only on fine-tuning.

Topics Covered:

- Why RAG instead of fine-tuning?
- Embeddings for retrieval
- Vector databases

- Similarity search
- Chunking strategies
- Hybrid search (keyword + vector)
- Prompt engineering basics
- Query rewriting techniques
- Building a simple chat interface

Minor Project:
Build a working RAG application:

- Upload document
- Generate embeddings
- Store in vector database
- Ask questions through chat interface
- Retrieve and generate accurate responses

Module 8 - "Bringing AI to Life"

Students move from building → improving and evaluating AI systems.

Topics Covered:

- Embedding optimization strategies
- RAG evaluation metrics
- Latency optimization
- Model benchmarking
- Research paper reading & discussion session
- Introduction to workflow automation (n8n)

Mini Projects:

- Improve RAG performance through chunking & embedding comparison
- Build a simple AI automation workflow (e.g., document upload → summary → notification)

PROJECT & REPORT (Weeks 8–10)

MAJOR PROJECT (Company Product Feature / Students Can come up with their own project)

Duration: 2 Weeks

Interns will work on:

- Real feature of company AI product
- Research-driven implementation
- Team Lead guidance
- Weekly review sessions

Final Report Structure

1. Introduction
2. Background & Related Work
3. System Architecture
4. Implementation
5. Results & Metrics
6. Performance Comparison
7. Challenges
8. Future Scope
9. Conclusion

Thank You & See you in classes!!

Contact: careers@thundertribes.com

WhatsApp: +91 8197261675