DataVerse Africa Internship Cohort 3.0

Data Science Track – 12 Weeks

Meeting Schedule Structure

- **First Mondays of Every being week of project:** Strategist session (domain context + project framing)
- **Tuesday**: Project assignment + foundational concepts + initial hands-on work (2 hours)
- **Thursday**: Advanced workshop + project development + progress review (2 hours)
- **Weekend**: Technical presentations every 3 weeks
- **Wednesday/Friday**: Self-study with coding exercises and independent project work
- **Time**: 8:00PM WAT

Week 0:

Icebreaker Project:

Weeks 1–3: Machine Learning Fundamentals

Project: *Nigerian Agricultural Yield Prediction System (can be changed)*

- **Week 1 – Vision & Foundation**
    - **Monday (Strategist):** *ML in African Agriculture/healthcare and Finance – Real-World Use Cases*
    - **Tuesday:** Introduction to ML workflow, supervised vs unsupervised. Live coding: scikit-learn basics (regression/classification).
    - **Thursday:** Advanced Workshop → Data preprocessing, feature engineering, and baseline model training.
    - **Saturday:** Presentation → "Baseline ML Model & Data Exploration."
    - **Self-study:** NumPy, Pandas, and EDA exercises.
    - **Deliverable:** *Baseline regression notebook + research brief*.
- **Week 2 – Design & Implementation**
    - **Tuesday:** Model building: regression, decision trees, evaluation metrics (MAE, RMSE).
    - **Thursday:** Advanced Workshop → Cross-validation, hyperparameter tuning, model comparison.

- **Saturday:** Presentation → “Prototype Models + Insights.”
- **Self-study:** ML model evaluation & scikit-learn practice.
- **Deliverable:** *Prototype ML models + evaluation report*.

- **Week 3 – Synthesis & Delivery**
  - **Tuesday:** Ensemble learning (Random Forest, XGBoost), interpretability (SHAP values).
  - **Thursday:** Advanced Workshop → Build end-to-end ML pipeline. Finalize results.
  - **Saturday:** Presentation → “Final ML Pipeline & Yield Prediction.”
  - **Self-study:** ML pipeline practice + documentation writing.
  - **Deliverable:** *Agricultural ML Blueprint + project presentation*.

Weeks 4–6: Deep Learning & Computer Vision

Project: *Crop Disease Detection System*

- **Week 4 – Vision & Foundation**
  - **Monday (Strategist):** *How AI & CV Improve African Food Security, Healthcare and Finance*
  - **Tuesday:** Neural networks basics, backpropagation, OpenCV image preprocessing.
  - **Thursday:** Advanced Workshop → Image dataset exploration, preprocessing pipeline design.
  - **Saturday:** Presentation → “Dataset Insights + Preprocessing Plan.”
  - **Self-study:** TensorFlow basics, image augmentation practice.
  - **Deliverable:** *Image preprocessing + dataset setup*.
- **Week 5 – Design & Implementation**
  - **Tuesday:** CNN architectures (LeNet, AlexNet, VGG). Implement baseline CNN.
  - **Thursday:** Advanced Workshop → Data augmentation, model training & validation.
  - **Saturday:** Presentation → “Baseline CNN & Model Metrics.”
  - **Self-study:** Keras/TensorFlow tutorials + CNN coding exercises.
  - **Deliverable:** *Trained CNN + evaluation report*.

- **Week 6 – Synthesis & Delivery**
    - **Tuesday:** Transfer learning (ResNet, EfficientNet), Grad-CAM for explainability.
    - **Thursday:** Advanced Workshop → Optimize and fine-tune CNN. Build demo.
    - **Saturday:** Presentation → "Final CV Model & Mobile Demo."
    - **Self-study:** Transfer learning experiments.
    - **Deliverable:** *Final disease detection model + demo presentation*.

Weeks 7–9: NLP, Generative AI & Cloud Deployment

Project: *African Agricultural Advisory Chatbot (Cloud-Hosted)*

- **Week 7 – Vision & Foundation**
    - **Monday (Strategist):** *NLP & Generative AI for African Languages & Rural Farmers*
    - **Tuesday:** Text preprocessing (tokenization, stemming, lemmatization). Tools: NLTK, spaCy.
    - **Thursday:** Advanced Workshop → Intent classification + sentiment analysis.
    - **Saturday:** Presentation → "NLP Preprocessing + Chatbot Design Plan."
    - **Self-study:** NLP text processing challenges in African languages.
    - **Deliverable:** *Chatbot intents + preprocessing pipeline*.
- **Week 8 – Design & Implementation**
    - **Tuesday:** Build baseline chatbot (rule-based + ML classification).
    - **Thursday:** Advanced Workshop → Dialogue flow design + integration with farmer queries.
    - **Saturday:** Presentation → "Working Chatbot Prototype."
    - **Self-study:** Chatbot frameworks (Rasa, simple transformers).
    - **Deliverable:** *Working chatbot prototype*.
- **Week 9 – Synthesis & Delivery**
    - **Tuesday:** Transformers (BERT, GPT), multilingual AI fine-tuning.
    - **Thursday:** Advanced Workshop → Deploy chatbot with Docker/API (Colab or cloud).
    - **Saturday:** Presentation → "Final GenAI Chatbot Deployed on Cloud."
    - **Self-study:** Cloud ML services (SageMaker/GCP/Azure basics).
    - **Deliverable:** *Cloud-deployed Generative AI chatbot + presentation*.

Weeks 10–12: Cross-Collaborative Capstone (With Data Analytics Track)

Theme: *Comprehensive African Agricultural Intelligence Platform*

- **Week 10 – Vision & Foundation**
  - **Monday (Strategist, Joint Session):** *Cross-Domain Integration for African Agriculture*
  - **Tuesday:** DS: Integrating ML, CV, NLP models into APIs. DA: Dashboard design for stakeholders.
  - **Thursday:** Joint Workshop → Unified project planning, architecture design.
  - **Saturday:** Presentation → "Capstone Plan + Initial Datasets."
  - **Deliverable:** *Capstone project plan + early prototype*.
- **Week 11 – Design & Implementation**
  - **Tuesday:** DS: Deploy ML pipelines on cloud. DA: Build dashboards linked to APIs.
  - **Thursday:** Joint Workshop → End-to-end integration testing.
  - **Saturday:** Presentation → "Integrated Prototype Demo."
  - **Deliverable:** *Integrated prototype + test results*.
- **Week 12 – Synthesis & Delivery**
  - **Tuesday:** Final integration, optimization, stakeholder-ready demos.
  - **Thursday:** Mock final presentation + feedback.
  - **Saturday (Graduation):** *Final Capstone Showcase & Panel Evaluation*.
  - **Deliverable:** *Final integrated platform + team presentation + documentation*.