



EURON

TECH

For Data Science Stack



1. Machine Learning (ML) Engineer Tech Stack

- **Programming Languages:** Python, R, Julia (occasionally)
 - **Libraries & Frameworks:**
 - **Basic ML:** NumPy, Pandas, Scikit-learn
 - **Advanced ML:** XGBoost, LightGBM, CatBoost
 - **Data Preprocessing:**
 - Feature Engineering: Scikit-learn, Pandas, Featuretools
 - Data Cleaning: OpenCV (for images), TextBlob (for NLP)
 - **Model Development:**
 - Regression, Classification, Clustering, Dimensionality Reduction
 - **Model Evaluation:**
 - Cross-validation, AUC-ROC, Precision-Recall, RMSE
 - **Big Data Tools** (for large-scale ML):
 - Spark MLlib, Dask, Ray
 - **Cloud & Deployment:**
 - AWS SageMaker, Google Vertex AI, Azure ML
 - **MLOps & CI/CD:**
 - Docker, Kubernetes, MLflow, Kubeflow, TensorFlow Serving, FastAPI
-

2. Deep Learning (DL) Engineer Tech Stack

- **Programming Languages:** Python, Julia
- **DL Libraries & Frameworks:**
 - TensorFlow, PyTorch, Keras, JAX
- **Deep Learning Techniques:**
 - CNN, RNN, LSTM, Transformers
- **Model Optimization:**
 - Mixed Precision Training (NVIDIA Apex)
 - Quantization (TensorRT, ONNX)
- **Hardware Acceleration:**
 - GPU: CUDA, cuDNN, TensorRT
 - TPU (for cloud-based training)
- **Cloud & Deployment:**

- AWS Inferentia, Google TPU Pods, Azure Machine Learning
 - **Containerization & CI/CD:**
 - TensorFlow Serving, TorchServe, FastAPI, Docker, Kubernetes
-

3. Natural Language Processing (NLP) Engineer Tech Stack

- **Programming Languages:** Python
 - **Text Processing Libraries:**
 - NLTK, Spacy, TextBlob, regex
 - **ML/DL Libraries for NLP:**
 - Hugging Face Transformers, BERT, GPT, T5, LLaMA
 - **Vectorization & Embeddings:**
 - TF-IDF, Word2Vec, FastText, GloVe, BERT embeddings
 - **Sequence Models:**
 - LSTM, GRU, Transformers (Attention Mechanisms)
 - **NLP Tasks & Tools:**
 - Named Entity Recognition (NER): Spacy, StanfordNLP
 - Sentiment Analysis: VADER, TextBlob
 - Summarization: T5, Pegasus
 - Chatbots: Rasa, Dialogflow
 - **Cloud-based NLP Services:**
 - AWS Comprehend, Google Cloud NLP, Azure Cognitive Services
 - **Deployment:**
 - FastAPI, Streamlit, Docker, Kubernetes, TensorFlow Serving
-

4. Computer Vision (CV) Engineer Tech Stack

- **Programming Languages:** Python, C++ (for efficiency)
- **Computer Vision Libraries:**
 - OpenCV, PIL (Pillow), Dlib
- **Deep Learning Frameworks for CV:**
 - TensorFlow, PyTorch

- **Pre-trained CV Models:**
 - YOLO (for object detection)
 - Detectron2 (Facebook AI)
 - EfficientNet, ResNet, MobileNet, Vision Transformers (ViTs)
 - **Specialized CV Tasks & Tools:**
 - Image Classification: EfficientNet, ResNet
 - Object Detection: YOLOv8, Faster R-CNN
 - Image Segmentation: U-Net, Mask R-CNN
 - Optical Character Recognition (OCR): Tesseract, EasyOCR
 - **Augmentation & Preprocessing:**
 - Albumentations, imgaug
 - **Cloud-based CV Services:**
 - AWS Rekognition, Google Vision AI, Azure Computer Vision
 - **Deployment:**
 - ONNX Runtime, TensorRT, NVIDIA DeepStream (for edge AI)
 - FastAPI, Streamlit, Docker, Kubernetes
-

5. Generative AI (Gen AI) Engineer Tech Stack

- **Programming Languages:** Python
- **Foundation Models & Architectures:**
 - Transformers: GPT-4, T5, BERT, LLaMA
 - Diffusion Models: Stable Diffusion, DALL·E, MidJourney
 - GANs (Generative Adversarial Networks): StyleGAN, CycleGAN
 - Large Language Models (LLMs): GPT-3.5/4, Claude, Mistral, Gemini
- **Libraries & Frameworks:**
 - Hugging Face Transformers, LangChain, llama_index
 - OpenAI API, Cohere, Anthropic Claude API
- **Prompt Engineering & Fine-tuning:**
 - LoRA (Low-Rank Adaptation)
 - RLHF (Reinforcement Learning from Human Feedback)
- **Cloud AI & APIs:**
 - OpenAI API, Google Vertex AI, AWS Bedrock
- **Vector Databases for RAG (Retrieval-Augmented Generation):**

- FAISS, Pinecone, ChromaDB, Weaviate
 - **MLOps for LLMs:**
 - LangChain Agents, Hugging Face Spaces, DeepSpeed
 - **Deployment & Optimization:**
 - FastAPI, Streamlit, Kubernetes, NVIDIA Triton Inference Server
 - **Security & Ethics:**
 - AI Red Teaming, Bias Detection, AI Governance (GDPR, AI Act)
 - **Generative AI Applications:**
 - Text-to-Image: Stable Diffusion, DALL·E
 - Text-to-Video: RunwayML, Synthesia
 - AI Avatars: HeyGen, DeepBrain AI
 - AI Code Generation: GitHub Copilot, Code Llama
-

Comparison Table for Quick Reference

Stack	Programming	ML/DL Frameworks	Special Tools	Cloud & Deployment
ML Engineer	Python, R	Scikit-learn, XGBoost	Pandas, SQL	AWS SageMaker, Azure ML
DL Engineer	Python, Julia	TensorFlow, PyTorch	CUDA, TensorRT	GCP TPU, AWS Inferentia
NLP Engineer	Python	Transformers, Spacy	Hugging Face, BERT	OpenAI API, AWS Comprehend
CV Engineer	Python, C++	OpenCV, YOLO, ViTs	Detectron2, Tesseract	AWS Rekognition, DeepStream
Gen AI Engineer	Python	LLMs, Diffusion Models	LangChain, RLHF	OpenAI API, Pinecone

Final Thoughts

- Each of these specializations requires a mix of: ✓ **Programming** (Python, SQL) and Bit or Math
- ✓ **ML/DL Frameworks** (TensorFlow, PyTorch, Hugging Face)
- ✓ **Specialized Tools** (YOLO for CV, GPT for NLP, FAISS for Gen AI)
- ✓ **Cloud & MLOps** (AWS, GCP, LangChain, Kubernetes)