

KEREM ATES

Karsiyaka, Izmir, Turkiye • +90 538 599 49 47 • clite.ka@gmail.com
[LinkedIn Profile](#) • [GitHub Profile](#) • [Personal Portfolio](#)

PROFILE

Linguistics graduate from Boğaziçi University, currently AI/LLM evaluator at Yandex; author of an empirical thesis comparing native Turkish speakers and large language models on syntactic acceptability judgments. Hands-on experience fine-tuning transformer models (DistilBERT, BERTurk, mBERT) across NLP tasks, building end-to-end Python data products, and developing commercial algorithmic trading tools — including a TradingView indicator suite sold at confluencearray.com. Native-level English (YDS 98.75 / 100).

EDUCATION

Boğaziçi University | *B.A. in Linguistics* Graduated: 2025

Senior Thesis: "Acceptability Judgments of Turkish Word Order: A Comparative Study of Native Speakers and Large Language Models."

- Designed and ran a structured Acceptability Judgment Task (AJT) with 73 native Turkish speakers; evaluated 5 LLMs (GPT-4o, Claude 3.7 Sonnet, Gemini 2.0 Flash, Grok 3, DeepSeek V3) on the same sentence sets across word order and animacy conditions.
- Found that LLMs successfully reproduce human preferences for canonical orders but show markedly reduced sensitivity to animacy and gradient acceptability, with significant variation across model architectures.

PROFESSIONAL EXPERIENCE

Yandex | *AI Data Annotator & Evaluator* February 2025 – Present | Remote

- Evaluate Yandex's LLM output quality through NLP-focused human feedback, applying linguistics background to assess consistency, accuracy, and language quality across multilingual content.
- Use Excel to organize, quality-check, and report on large evaluation datasets.

Self-employed | *Independent Software Developer* June 2024 – Present | Remote

- Design and ship end-to-end software projects independently — spanning NLP/ML applications (NER, stance detection, cross-lingual transfer, predictive modelling) and production trading tools across TradingView, Streamlit, and MetaTrader 5.
- Handle the full development lifecycle independently — from Python prototyping and backtesting through to live deployment and commercial distribution.

Self-employed | *Quantitative Trader* January 2021 – Present | Remote

- Develop and trade systematic strategies across forex and CFD markets, managing the full research-to-execution workflow — hypothesis testing, multi-year backtesting, and live deployment.
- Apply structured risk management: defined R:R targets, position sizing rules, and trade logging for ongoing performance review.

PROJECTS

Named Entity Recognition with DistilBERT | *Python, HuggingFace, DistilBERT, PyTorch, Gradio*

github.com/cliteka-cell/ner-wikiann • [Live Demo](#)

- Fine-tuned DistilBERT on WikiANN English (20k sentences) for PER/ORG/LOC tagging; achieved F1 = 0.82.
- Manual error analysis on 80k test tokens showed 74% were misclassifications (mainly LOC/ORG ambiguity) rather than missed detections; deployed as interactive HuggingFace Space.

Stance Detection with DistilBERT | *Python, HuggingFace, DistilBERT, PyTorch, Gradio*

github.com/cliteka-cell/stance-detection • [Live Demo](#)

- Fine-tuned DistilBERT on the IBM Argument Quality dataset (30k arguments) to classify debate arguments as PRO/CON; achieved F1 = 0.90 on a balanced dataset.
- Analyzed stance ambiguity and model handling of complex argumentative language; deployed live demo on HuggingFace Spaces.

Cross-Lingual Transfer: Turkish Sentiment Analysis | *Python, HuggingFace, mBERT, PyTorch, Gradio*

github.com/cliteka-cell/turkish-sentiment-crosslingual • [Live Demo](#)

- Demonstrated cross-lingual failure: an English sentiment model on Turkish scored 41.7% accuracy (worse than random); fine-tuned mBERT on 4,486 Turkish restaurant reviews, recovering to 84.4% (+42.6% gain).
- Analyzed Turkish morphology, tokenization failure modes, and within-language domain mismatch.

ARMAS — Algorithmic Expert Advisor for MetaTrader 5 | *Python, MQL5, Optuna*

github.com/cliteka-cell/ARMAS_Project_Full

- Mean-reversion Expert Advisor for MetaTrader 5 using an EWM band strategy; iterated from Python backtest to production EA across 18 versions.
- Implements RSI/ATR/slope signal filters, Optuna parameter tuning, and ML training data generated from live trade logs.

Confluence Array — TradingView Indicator Suite | *Pine Script v6, TradingView*

confluencearray.com • [TradingView Script](#)

- Commercial multi-mode indicator that auto-detects trading timeframe (intraday / hourly / macro) and loads the relevant levels — VWAP, market profile, FVGs, ARMAS bands, CVD divergence, open interest signal matrix, and a scoring engine ranking zones by confluence strength.
- Sold commercially at confluencearray.com; used daily in live trading.

Additional Projects

- **ARMAS RL Agent** — MaskablePPO reinforcement learning; documented negative result. github.com/cliteka-cell/armas-rl
- **Trading Journal** — Streamlit app with MT5 import, equity curve, and PDF export. trading-journal-preview.streamlit.app
- **ALES Oracle** — Gemini 2.5 Flash AI exam prep tool. ales-oracle.streamlit.app
- **İzmir Acil** — GPS-based emergency services directory with 4,600+ listings. izmir-acil.streamlit.app
- **House Price Prediction** — XGBoost regression model; $R^2 = 0.944$.
- **E-Commerce Customer Behaviour** — RFM segmentation across 540k transactions.
- **Open-Source TradingView Indicators** — nPOC Daily, Session CVD Divergence, Auto Anchored VP, SFP Detector. github.com/cliteka-cell

TECHNICAL SKILLS

Programming: Python, SQL, R, C#, MQL5, Pine Script

Machine Learning: Scikit-learn, XGBoost, Optuna, Feature Engineering, Model Evaluation

NLP & Transformers: HuggingFace, PyTorch, DistilBERT, mBERT, Fine-tuning, NER, Token Classification, Stance Detection, Cross-Lingual Transfer

AI / LLMs: LLM Evaluation, Prompt Design, Gemini API, Data Annotation & Quality Control

Data Analysis: Pandas, Matplotlib, Seaborn, Excel, EDA & Statistical Analysis, Geospatial Analysis (Folium, Geopy)

Web & Deployment: Streamlit, Gradio, HuggingFace Spaces, SQLite, Web Scraping

Finance & Trading: Algorithmic Trading, Backtesting, Order Flow Analysis, Risk Management, MetaTrader 5

LANGUAGES & CERTIFICATIONS

- **English** — Native-level proficiency (C2)
- **YDS English Certification** — 98.75 / 100 (Level A, highest tier) — April 2026, ÖSYM
- **ALES Quantitative** — 70.627 — Numerical reasoning university entrance exam, ÖSYM
- **Spanish / German / French** — Beginner (A1)

ADDITIONAL INFORMATION

Interests: AI and data science, computational linguistics, quantitative finance, culture and arts publishing