

# RoastFormer: Transformer-Based Time-Series Generation for Coffee Roast Profiles

**1. Project Title and Overview:** Coffee roasters develop roast profiles through trial-and-error: each new coffee requires 10-20 experimental roasts (~15 minutes each plus cupping time) to build a temperature-over-time curve from scratch. Modern logging software (Artisan, Cropster) captures and replicates successful profiles, but no predictive starting point exists; roasters must experiment from zero for every coffee, manipulating heat based on sight, smell, and sound (first crack!), then evaluating through cupping protocols. **RoastFormer** investigates whether transformers designed for discrete text can generate physically plausible continuous time-series profiles conditioned on bean characteristics (density, origin, processing) and desired outcomes (light/medium/dark). Using physics-based synthetic data validated against real Artisan CSV exports, I will implement Phuong & Hutter's DTransformer, treating roasting as autoregressive sequence modeling where temperature readings are generated token-by-token analogous to GPT's text generation.

**2. Connection to Course Material:** This project applies transformer architectures to domain-specific continuous time-series data, exploring how components designed for discrete text tokens transfer to physical processes. Using Formal Algorithms as the implementation foundation, I investigate key technical tradeoffs: positional encoding strategies for temporal data (learned vs. RoPE – my paper presentation), attention pattern interpretability (do heads learn roasting phases?), and model generalization from synthetic to real-world data. This demonstrates practical understanding of transformer fundamentals while evaluating which architectural choices matter beyond language modeling.

**3. Problem Statement and Goals:** Coffee roasters lack predictive tools for creating initial roast profiles, relying entirely on field experience. This project addresses whether transformers can generate physically plausible continuous temperature curves for a novel coffee given only its characteristics (density, origin, target roast level). Understanding which architectural components enable this adaptation matters because the findings generalize to any continuous time-series generation problem: industrial process control, climate forecasting, and financial modeling all face similar challenges translating discrete-token architectures to continuous temporal data. Goals include: (1) Adapt decoder-only transformer for continuous temperature sequences; (2) compare positional encoding methods on generation quality; (3) analyze if attention learns roasting phases without supervision; (4) validate synthetic profiles against real Artisan exports; (5) identify which components matter through ablations. Overall success looks like generating profiles matching real Artisan characteristics (3-second intervals, 10-15 minutes, 70-450°F, smooth transitions) with <5°F average error.

**4. Timeline:** **Oct 20-25:** Project definition, proposal submission. **Oct 27-Nov 1:** Generate 10K synthetic profiles validated against 2 real Artisan exports, establish baselines. **Nov 3-8:** Implement DTransformer with 3 positional encoding variants. **Nov 10-15:** Run ablations, analyze attention, create visualizations. **Nov 17-20:** Finalize presentation (real vs. generated comparison), Model Card, pseudocode.

**5. Success Metrics / Assessment Methods:** Model accuracy will be based on Mean Absolute Error <5°F for temperature predictions, meeting practical roasting tolerance. Qualitative analysis to include visualizing attention patterns to verify if heads specialize on roasting phases without supervision; confirm dense beans generate 20-25% slower heating as physics predicts; compare generated curves against real Artisan exports for structural realism (smooth transitions, 3-second sampling, realistic heating rates 20-100°F/min). Physical plausibility will verify 100% monotonic temperature increase, >95% bounded heating rates, >90% reach targets, ensuring outputs obey roasting physics rather than producing impossible curves.

**6. Value Proposition:** For roasters, this project demonstrates proof-of-concept for data-driven starting profiles that could reduce experimental iterations and coffee waste. This tool would be particularly valuable for junior roasters lacking extensive cupping experience, providing educated starting points while they develop sensory skills. Real-world deployment would require training on actual roast data and should augment, not replace, roaster expertise and sensory evaluation. As for practitioners, reusable pseudocode and synthetic data framework validated against real Artisan exports provide a template for transformer-based time-series generation across domains (industrial process control, climate modeling, financial forecasting), demonstrating how to adapt NLP architectures to continuous temporal data.

