Automating Model Training & Ensuring Ethical Model Development

AI Masters Capstone Project - Presentation 3

Jonathan Agustin

November 2024

What We'll Cover Today

- Automating model training: efficient pipelines, hyperparameter optimization
- Model versioning & experiment tracking: ensuring reproducibility
- Transparent & explainable AI: building user trust
- Fairness in AI models: evaluating & mitigating biases

From raw code to final model deployments, ethics and efficiency guide our steps.

Automating Model Training

- Reduce manual iteration: automate data loading, training, validation steps
- Integrate hyperparameter search methods (grid, random, Bayesian) seamlessly
- Create reproducible pipelines that encourage systematic improvement

Automation frees us to concentrate on meaningful insights and better outcomes.

Hyperparameter Optimization

- Use automated search strategies (Bayesian, genetic algorithms) to tune hyperparams
- Track performance at each trial, building knowledge over time
- Achieve superior models with fewer manual experiments

Smart HPO transforms trial-and-error into a guided search for excellence.

Model Versioning & Experiment Tracking

- Tag each model, dataset, and parameter set for easy reference
- Log metrics, outcomes, and conditions for every run (e.g., MLflow)
- Instantly reproduce past experiments, fostering accountability and trust

A well-documented history of experimentation ensures transparency and continual learning.

Transparent & Explainable AI

- Apply tools like SHAP, LIME to clarify model decisions
- Visualize feature importances & decision pathways
- Enhance stakeholder trust by revealing the reasoning behind predictions

When ML decisions are understandable, we foster trust, alignment, and responsible use.

Fairness in AI Models

- Check metrics like demographic parity, equality of odds for different groups
- Apply interventions (re-sampling, constraints) to reduce unfair outcomes
- Continuously monitor fairness, ensuring changes don't reintroduce bias

Fair models reflect our commitment to justice, equity, and ethical innovation.

A Holistic Model Development Approach

- Combine automation, tracking, explainability, and fairness measures seamlessly
- Document the entire workflow, from preprocessing to final model
- Align performance with ethical principles for responsible ML innovation

The result: models you can trust, understand, and be proud to deploy.

Next Steps

- Next Presentation: Deployment Automation & Ethical Deployment
- Ensuring secure, compliant, user-centered model releases

From fair data to fair models, we're shaping AI that serves humanity responsibly.

Temporary page!

LaTeX was unable to guess the total number of pages correctly. As there was some unprocessed data that should have been added to the final page this extra page has been added to receive it.

If you rerun the document (without altering it) this surplus page will go away, becau LaTeX now knows how many pages to expect for this document.