# Research Statement - Joseph Paul

I am a Ph.D. candidate in Economics at Heriot-Watt University, Edinburgh. My research develops practical methods for evaluating, improving, and deploying machine learning algorithms in complex decision-making contexts, particularly where reliability, fairness, and real-world impact are important. Sitting at the intersection of econometrics, statistics, and machine learning, my work aims to bridge the gap between theoretical advancements in algorithmic design and the challenges faced in practical application. My current research can be broadly categorized into two main strands: (1) the evaluation and improvement of algorithmic policies based on their welfare and performance impacts, and (2) the development of reliable uncertainty quantification techniques for predictive settings, especially non-standard ones.

## Evaluating and Improving Algorithmic Policies

While algorithms show significant promise in domains like public policy, healthcare, and finance, their evaluation frequently centres on predictive accuracy, potentially overlooking broader impacts. Additionally, theoretical frameworks for optimal policy learning, such as Empirical Welfare Maximization, often provide performance guarantees (e.g., regret bounds) under specific assumptions about the data generating process or in large-sample regimes. Bridging the gap between these theoretical results and practical deployment requires addressing challenges related to finite-sample performance, complex real-world constraints, and ensuring desirable outcomes beyond simple prediction metrics. My work addresses this gap by developing a pragmatic framework for assessing algorithmic decision-making systems.

In “Testing Algorithmic Welfare Improvability” (working paper) I propose a statistical testing procedure for evaluating algorithmic improvability across multiple objectives. This framework allows an analyst to test if a proposed or existing algorithm can be improved upon by another algorithm from a potentially complex, user-defined class (e.g., deep neural networks), without sacrificing performance on other specified objectives such as fairness, profit, or accuracy, and while respecting operational constraints. The methodology leverages doubly robust estimation of conditional average treatment effects for welfare evaluation and employs bootstrapping and sample-splitting, accommodating a wide range of algorithm types and utility functions.

This research is motivated by the need for practical evaluation tools that enable policy-makers, regulators, and developers to assess whether these systems meet their intended objectives and to identify opportunities for improvement. It provides a means to audit algorithms and guide their responsible deployment, particularly in areas like disparate impact analysis.

## Robust Uncertainty Quantification

Effective decision-making often requires not just point predictions, but also a reliable assessment of uncertainty. Conformal prediction offers a powerful, distribution-free approach to constructing prediction intervals with guaranteed coverage. However, existing methods frequently rely on a single predictive model, which may be suboptimal, as well as assuming exchangeable data, limiting applicability in non-standard settings like time series forecasting.

In “Ensemble Methods for Conformal Prediction” (working paper), I develop a method for ensembling conformal prediction sets from multiple models under the assumption of exchangeability. By aggregating inclusion decisions based on a threshold derived from concentration inequalities, the ensemble set provably achieves the desired overall coverage level, providing a principled way to leverage the prediction sets of several models.

Building on this, “Online Adaptive Ensembled Conformal Prediction for Time Series” (working paper) tackles the challenges of non-exchangeable data and potential distribution shifts, common in time series data. I propose an Online Adaptive Ensembled Conformal Prediction algorithm that dynamically adjusts both the target coverage level of individual models and their contribution to the ensemble based on their recent performance (balancing coverage and interval size). This online learning approach allows the ensemble to adapt to changing data dynamics while aiming to maintain marginal coverage guarantees over time. The method has provable finite coverage under mild distribution shifts, which I further explore in Monte Carlo simulations, and apply to the task of forecasting natural gas consumption.

Overall, my research seeks to develop statistically sound, practically relevant tools that enhance the transparency, reliability, and beneficial impact of algorithmic systems in real-world applications.