VISHAL CHAUDHARY · DATA ANALYST · DUBLIN, IRELAND

Causal Machine Learning & Research · Academic Project

Causality in Video Recommendations

SDID causal validation achieving 54.2% precision improvement over DiD on 3.13 million Kuaishou interactions.

3.13M	94.36%	+54.2%	7
User Interactions	Best AUC Score	SDID Precision Gain	ML Models Tested

TOOLS & TECHNOLOGIES

Python	SDID	XGBoost	Causal Inference	Pandas	Scikit-learn
Difference-in-Differences	R				

Email vishal.ch1401@gmail.com	LinkedIn linkedin.com/in/vishal111	GitHub github.com/chaudhary521	Location Dublin, Ireland

PROBLEM STATEMENT

Standard recommendation systems optimise for engagement correlation — they surface content users are likely to engage with. But a video a user would have found organically is not a meaningful recommendation. This MSc research asked: how much of the apparent effectiveness of recommendation systems is genuine causal lift versus pre-existing user preference? And can machine learning models be reliably evaluated for causal rather than correlational performance?

DATASET

KuaiRand-1K dataset: 3.13 million user-video interaction records from the Kuaishou short-video platform. The dataset includes watch time, like and share behaviour, and critically — randomised exposure logs that enable genuine causal identification. 1,000 users and 7,583 unique videos. The randomised exposure component is rare in industry datasets and what makes causal analysis tractable.

APPROACH & METHODOLOGY

Seven ML models were benchmarked on both standard accuracy metrics and causal effectiveness metrics. Synthetic Difference-in-Differences (SDID) was implemented as the primary causal framework — it constructs synthetic control groups from pre-exposure trends to isolate the true causal effect of a recommendation, going beyond standard DiD which assumes parallel trends. XGBoost was found to be the strongest standard model. The SDID framework was compared against DiD and naive correlation metrics.

KEY TECHNICAL HIGHLIGHTS

› Implemented the full SDID framework from scratch in Python, constructing synthetic controls from pre-exposure engagement trends.

› SDID achieved 54.2% higher causal precision than standard Difference-in-Differences methodology.

› XGBoost achieved the best overall AUC of 94.36% among all seven benchmarked models.

› Demonstrated that correlation-optimised models systematically overestimate recommendation effectiveness by 31–54%.

› KuaiRand-1K's randomised exposure logs enabled genuine causal identification without instrumental variable assumptions.

› Research submitted and assessed as part of the MSc Data Analytics programme at NCI Dublin.

KEY INSIGHTS & RESULTS

Standard ML models overestimate recommendation effectiveness by 31–54% due to selection bias — users who engage with recommended content frequently would have found it without the recommendation. The SDID framework isolates true incremental lift, revealing which recommendations genuinely caused additional engagement. XGBoost's strong AUC on standard metrics did not translate to proportionally strong causal lift, illustrating the gap between predictive and causal performance.

BUSINESS IMPACT

Provides a rigorous, practically implementable framework for evaluating recommendation system causal effectiveness beyond A/B testing. Applicable to any platform where distinguishing causal impact from correlation is commercially significant: streaming, e-commerce, news, and social media. Reducing wasted recommendation exposure (recommendations that would not have been needed) improves true ROI and reduces the risk of filter-bubble effects.

This case study is part of Vishal Chaudhary's data analytics portfolio. For more projects and contact details visit: github.com/chaudhary521