**Title of Talk/Workshop**

Hands-on Intro to developing Explainability for Recommendation Systems

**Submission type \***

90 minutes workshop

**Abstract**:

Over the last decade, the commercial use of recommendation engines/systems by business has grown substantially, enabling the flexible and accurate recommendation of items/services to users. Examples of popular recommenders include (to name a few) movies, videos and books recommendation engines offered by Netflix, Youtube and Amazon respectively.

In general, most recommender systems are typically “black-box” algorithms trained to provide inference of relevant items to users using techniques such as collaborative or content-based filtering models or hybrid models. The algorithms used in these systems are broadly opaque, thus making the predicted recommendations lack full interpretability/explainability. Making recommenders explainable is very essential, as they try to provide transparency and address the question of why were particular items recommended by the engine to users/system designers.

Over the last few years there has been a growing area of research and development in explainable recommendation systems. Explainable recommendations systems are generally classified as Post-hoc (i.e. explainability is done post-recommendation) or Intrinsic (explainability is integrated into the recommender model) approaches. This workshop will provide a hands-on implementation of some of these approaches.

During this workshop a hands-on walk-through of each implemented approach will be demonstrated interactively in notebooks. Each approach will consist of the following steps: 1) Exploratory data analysis and pre-processing of the case study dataset used in this workshop, 2) implementation, training and validation of explainable recommendation system models, 3) prediction of recommendations and finally 4) analysis/evaluation of recommendation inference results and explainability metrics.

Specifically, the workshop will outline the following implementations:

1. Exploratory data analysis
2. Post-hoc approaches:
   1. Recommendations computed via collaborative filtering and explained using Association Rules
   2. Recommendations computed via Factorization Machines (FM) and explained using Locally Interpretable Model Agnostic Explanations (LIME)
3. Intrinsic approaches:
   1. Recommendations computed via Matrix Factorization (MF) and augmented/explained with neighbourhood-based explanation.
   2. Recommendations computed via Alternating Least Square (ALS) and augmented/explained with item-style explanation.
4. Brief introduction to explainable recommendations using Knowledge Graph-based models
5. Review of the findings and conclusions

The workshop will use the open source 100K Movielens dataset provided https://grouplens.org/datasets/movielens/.

All of the python libraries/packages (requirements.txt) and reference papers used in this workshop are open source and can be sourced from this Github repo: https://github.com/aidowu1/Hands-on-Intro-to-building-Explainability-for-RecSys

Can your talk be classified into any of the Themes below? \*

* Tools, frameworks & libraries

Audience Level \*

* Those who are at an intermediate level of understanding and experience with ML