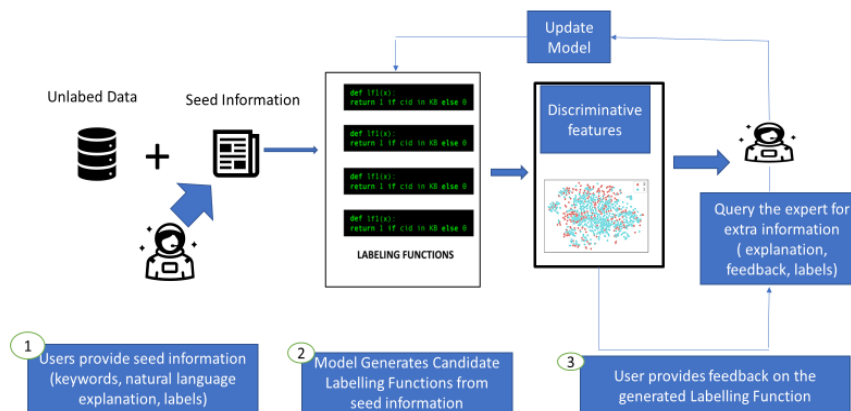




## Motivation

- Modern machine learning methods especially deep learning methods have achieved a state-of-art performance across several applications
- The performance of these models is explained by extremely large training datasets and expensive computational resource
- Their applications remains limited in low-resource environments
- Most of them are computational in nature and eliminate the user expertise in the analysis loop

## Weakly Supervised Learning and Subset Selection



## Solutions being studied

- Can we train models on smaller subsets rather than full-training set with minimal loss in performance?
- Can we leverage existing domain knowledge for training these models
- Can we Involve the user in the analysis loop through visualisations

### Weakly Supervised Learning

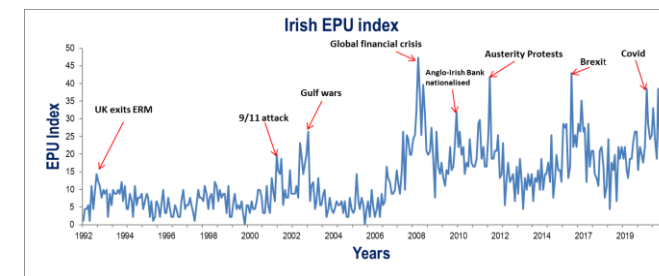
- Users express the domain knowledge as seed information to support model training
- Weak supervision uses the provided seed information to generate labelling functions that are used in labelling data which can be used in training other models

### Subset Selection

- We train models on carefully selected subsets
- The selected subsets are selected to minimise redundancy and approximate the full-training set properties as close as possible

## Applications

Analysis of documents posted in media outlets and social networks and their impact on society, both in historical and current data sets  
We used weak Supervision to generate an Economic Policy Uncertainty Index, an important index in monitoring the economy



## Works Done and in Progress

- Weak Supervision in the Analysis of News: Applications to Economic Policy Uncertainty (Done)
- Understanding High-Dimensional Spaces through visual means employing multidimensional Projections (Done)
- Interactive Weakly Supervised Learning: A Visual Analytics Approach (InProgress)
- Visual Analysis of New and Emerging Subset Selection Algorithms (InProgress)

## Contact

Email:12022601@uamail.ucc.ie