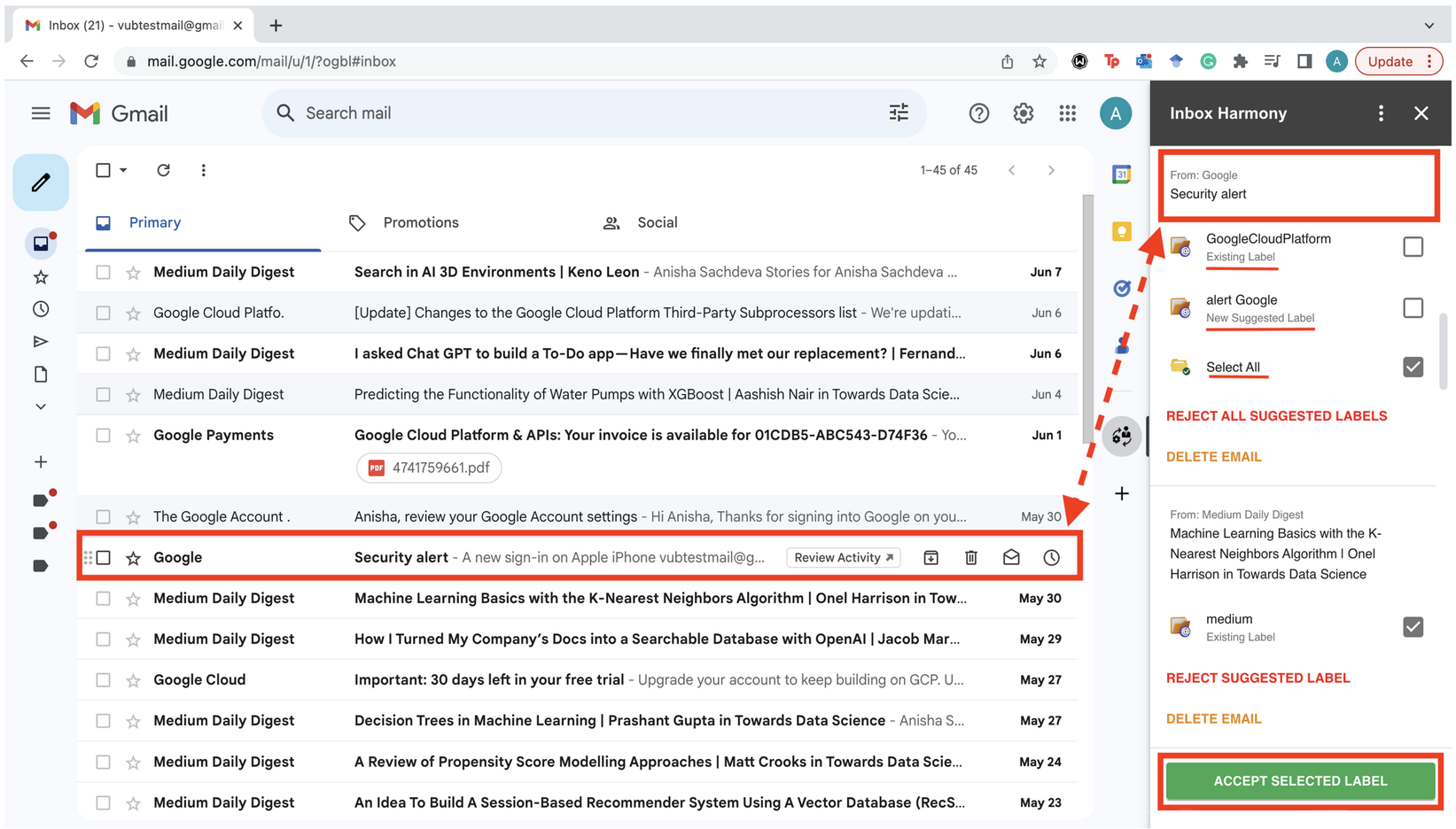
Inbox Harmony: A Recommendation System to Manage and Organise Emails Based on PIM Principles

Anisha Sachdeva and Beat Signer
(sachdeva.anisha@yahoo.com, bsigner@vub.be)

VRIJE UNIVERSITEIT BRUSSEL



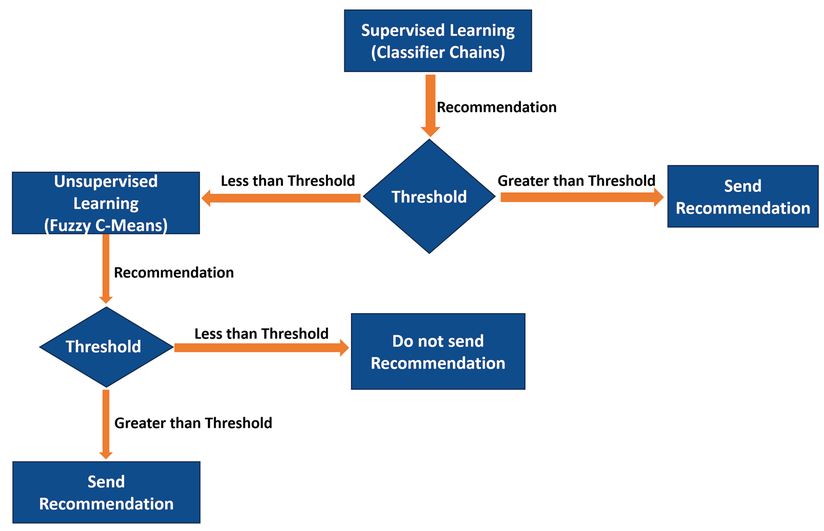
Research Question

Email is one of the primary forms of communication and to share information over the Internet. The increasing dependency on emails as a primary communication medium has led to a growing need for *efficient management strategies*, resulting in the following questions:

- What is *lacking in existing email clients*, *restricting users from easily managing their mailbox?*
- Can we use *machine learning algorithms to recommend actions* on similar emails, helping users to *automatically manage* their mailbox *by keeping the human in the loop?*

Contributions

We conducted a *survey* to get a deeper understanding of *current email management practices*. Further, we developed and evaluated a *recommendation system called Inbox Harmony* for email management. Inbox Harmony is an add-on available for Gmail with personalised recommendations. In scenarios where labelled data is available (user has already created some folders), the system uses *Classifier Chains*, a supervised machine learning model. However, if a user has not yet defined any folders, the system finds similarities within the unlabelled emails in a user's inbox, groups them, provides a name for the group using the *Latent Dirichlet Allocation* topic modelling approach and then present the newly suggested folders for the emails. In such scenarios where labelled data is not available, the system uses *Fuzzy C-Means*, an unsupervised machine learning model. A *hybrid deployment approach* is utilised by training both these models on an individual's Gmail account and providing personalised yet accurate machine learning-based recommendations to move a user's emails from their inbox to other folders. In contrast to existing rule-based filtering where an email can only be assigned to a signle folder, *Inbox Harmony can suggest multiple folders for a give email*.



Future Work

In the future we plan to address *platform compatibility* as well investigate *improved machine learning algorithms* addressing the *class imbalance problem* and profiting from *real-world datasets*.

Keywords : Email Management, NLP, Multi-Label Classification