Recommending similar items in large-scale online marketplaces

1. Please summarize and try to assess the approach described in this paper http://bit.ly/2oYesuo. What is the motivation? What are the key ideas? How does the approach differ from collaborative filtering or a "naive information retrieval system". What are possible shortcomings of the approach? Can you think of possible extensions? Please prepare a short presentation to 2-3 people that have a machine learning background.

2. In this dataset http://bit.ly/2qV472p you will find a sample of listings taken from the south african olx site as well as a list of popular recent keywords on olx.co.za
    a. Implement a simple approach to cluster the listings data into similar items clusters that can be used in this way: If the items A and B are part of the same cluster it is reasonable to recommend B to a user that is viewing or has viewed A. Feel free to use the category information that sellers have used to classify their listings in any way.
    b. How do you evaluate the quality of your results? How does it compare to a naive approach that takes random listings from the same category?
    c. What are possible shortcomings and extensions of your implementation? How are newly listed (unseen) listings assigned to your clusters?
    d. [Optional] If you like the clustering approach (the offline part) described in the paper feel free to implement a MVP using both datasets. Go ahead and simplify the task if you want or only use a subset of the data.
    e. Please submit your code in a GitHub repository and share it 24h before the presentation