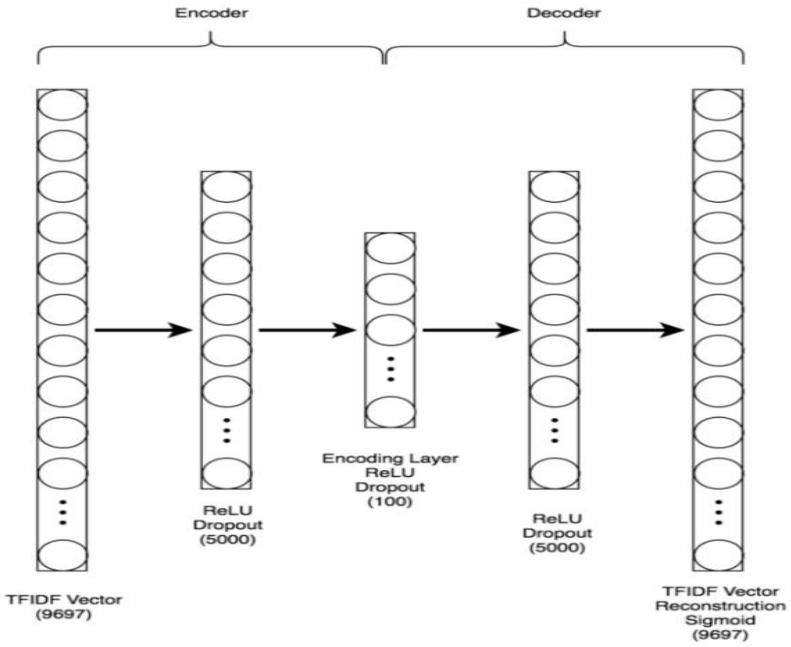
Hybrid based Recommended System

We will be using deep learning models to create a hybrid recommender system that uses both content and collaborative data. In this approach we would be using content and collaborative data separately at first, then combine the result that produces the best result from both. To create this model, we will take the results of an autoencoder which will learn content-based movie embeddings from tag data and a deep entity embedding neural network which will learns collaborative-based movie embeddings from ratings data.

Movie Embeddings from Content Based:

1. The tag documents into a Term Frequency Inverse Document Frequency (TF-IDF) representation.
2. In TF-IDF space, each dimension represents how important a certain word is to a movie. This representation is less than ideal because the encoding is fragmented and scattered across multiple dimensions. We would compress the TF-IDF data into a lower dimensional space where concepts are consolidated into shared dimensions, this we do using autoencoders.
3. Using the content-based movie embeddings learned by the autoencoder, the top N most similar movies can defined by cosine similarity.

Architecture is shown below:



Layer dimensions change depend on input

Movie Embeddings from Collaborative Data:

1. A neural network is used to find movie and user embeddings. In this architecture, a user and a movie embedding matrix are randomly initialized and subsequently learned using SVD.
   Once the network is satisfactorily trained, we will have movie and user embeddings ready to use for a variety of practical tasks. Using the collaborative-based movie embeddings learned and user embeddings by this neural net, the top n predicted rating of the movie for a certain user are given.