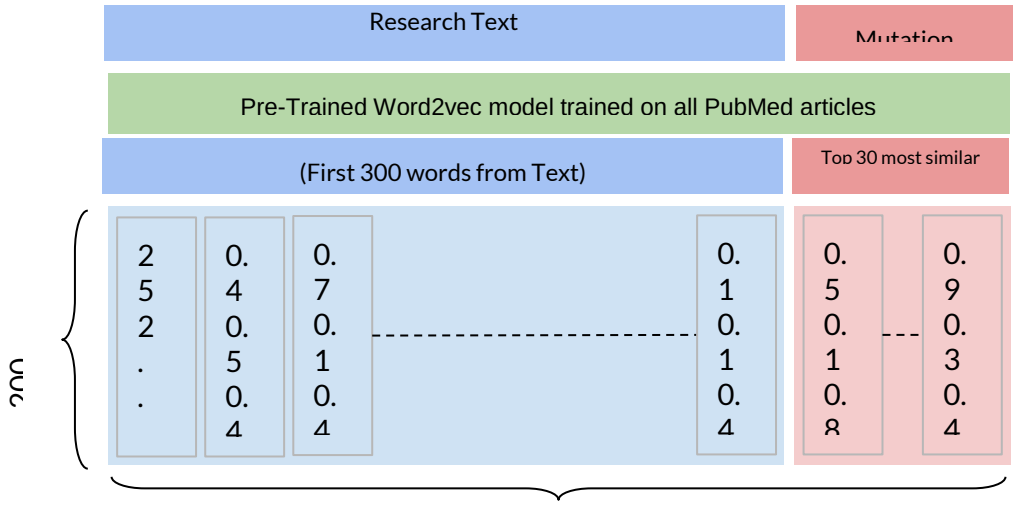


Figure 1: Data to sequences

Results

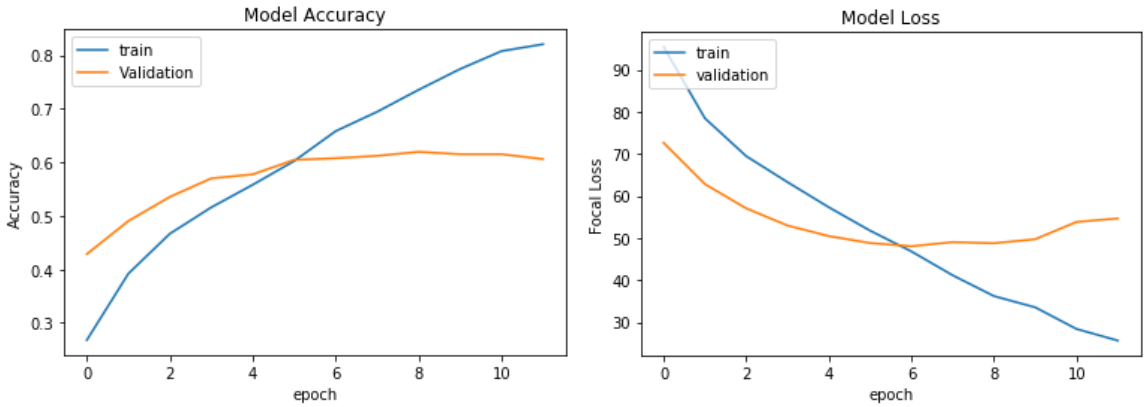


Figure 2: Bidirectional GRU with attention

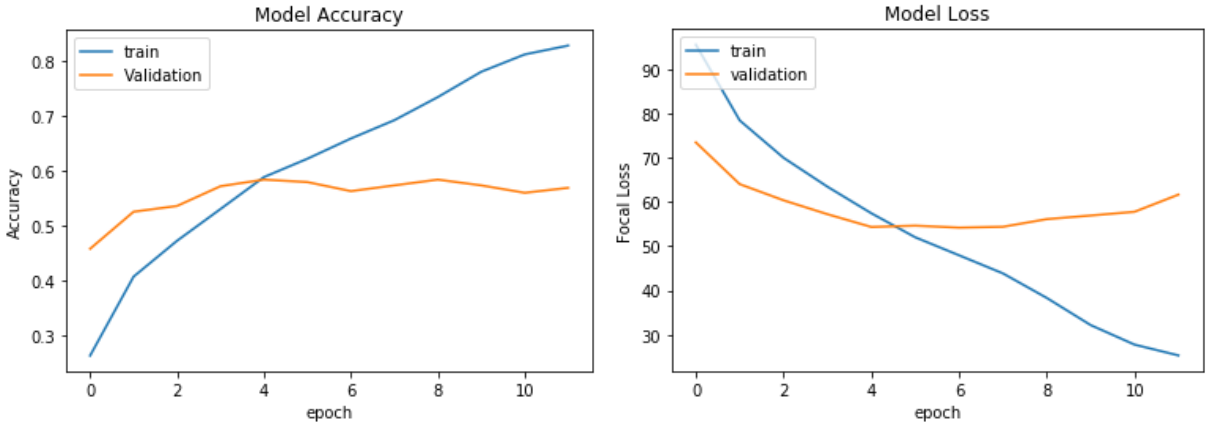


Figure 3: Bidirectional GRU

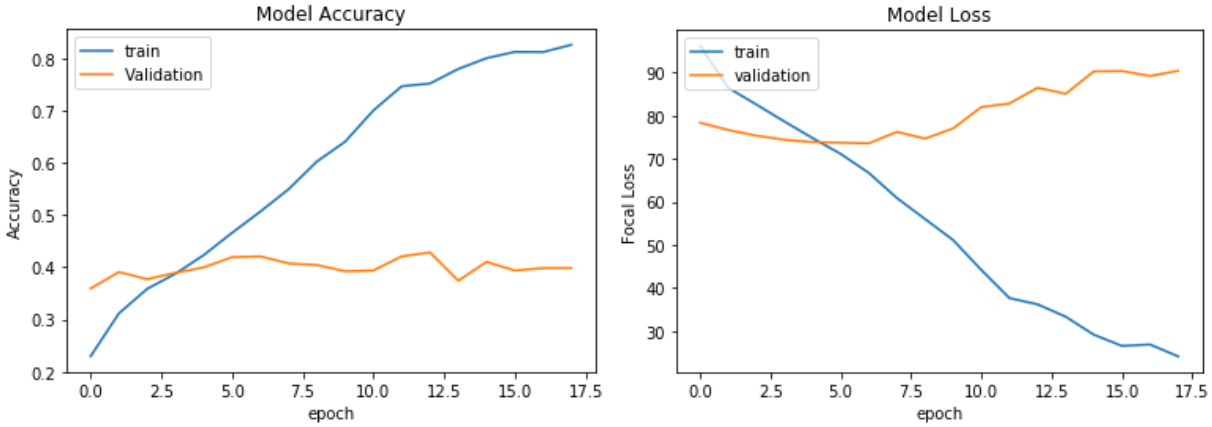


Figure 4: Simple GRU

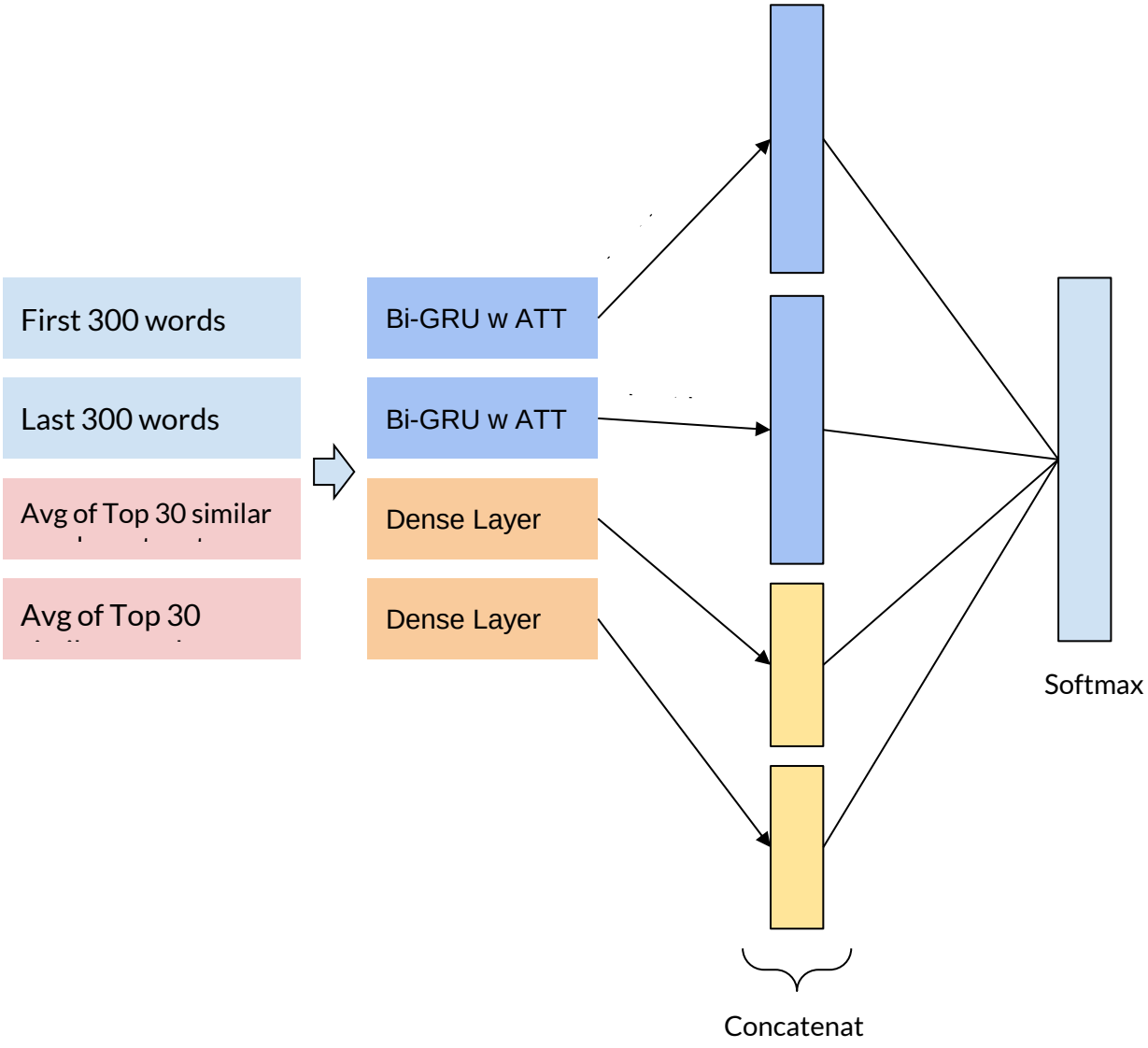
Sr. No	Model Description	Private Leaderboard Score	Public Leaderboard Score
1	Top in public leaderboard (Li-Der)	2.64525	0.10909
2	Top in private leaderboard (ilmirashaim)	2.03027	1.47447
3	Simple LSTM + Trainable Embeddings + No Batch Norm	2.58223	1.79227
4	Bi-LSTM + focal loss + Trainable Embeddings	**2.46368**	1.32249
5	Bi-GRU + focal loss + Trainable Embeddings	2.52356	1.30881
6	Bi-GRU with Attention + focal loss + Trainable Embeddings	2.51774	**1.29545**

Table 1. Some Kaggle Submissions

- We have tried to tackle the problem of class imbalance by using class weights and **focal loss**[1]

$$\mathrm{FL}(p_t) = -(1-p_t)^{\gamma}\log(p_t).$$

Future Experiments



[1] **"Focal Loss for Dense Object Detection." 7 Aug. 2017, https://arxiv.org/abs/1708.02002. Accessed 21 May. 2018.**