Need to label figure 2

1. The amount of data available is shown below. I will add the table tomorrow

|  |  |
| --- | --- |
| **Domain** | **Data** |
| Toys and Games(toys) | 2,252,771 reviews |
| Video Games(vid) | 1,324,753 reviews |
| Automotive(aut) | 1,373,768 reviews |
| Home & Kitchen(hnk) | 4,253,926 reviews |

1. Should we have a separate section for data preprocessing?
2. In section 4.3, that the transfer loss was approximately correlated with the domain similarity ie transfer loss was lower when transferring from a more similar domain.- *we can say that this inverse correlation of transfer loss with domain similarity is very intuitive.*

|  |  |  |  |
| --- | --- | --- | --- |
| **Active Transfer** | **Vocab** | **Pred\_prob** | **Counts** |
|  | vid | Pos probabilities > 0.9 | 9348 |
|  | vid+aut | Pos probabilities > 0.9 | 2739 |
| Y | vid+aut | Pos probabilities > 0.9 | 3256 |
|  | vid | Pos probabilities between 0.8 and 0.9 | 2937 |
|  | vid+aut | Pos probabilities between 0.8 and 0.9 | 1207 |
| Y | vid+aut | Pos probabilities between 0.8 and 0.9 | 841 |

**False positives table**

|  |  |  |  |
| --- | --- | --- | --- |
| **Active transfer** | **Vocab** | **Pred\_prob** | **Counts** |
|  | vid | Neg probabilities > 0.9 | 6146 |
|  | vid+aut | Neg probabilities > 0.9 | 1619 |
| Y | vid+aut | Neg probabilities > 0.9 | 1270 |
|  | vid | Neg probabilities between 0.8 and 0.9 | 4195 |
|  | vid+aut | Neg probabilities between 0.8 and 0.9 | 1146 |
| Y | vid+aut | Neg probabilities between 0.8 and 0.9 | 491 |

False negatives table