**SUMMARY OF FINDINGS REGARDING EXPERIMENTS WITH OM DATA AND**

**RNN LANGUAGE MODEL.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **SL** | **Experiment Name** | **Notebook & Dataset URL** | **Samples Used** | **Result** | |
| 01 | Exp 1 Jan23 - Regular Dataset - Tokenization 1  6 OM Combined | Notebook:  <https://tinyurl.com/expr1-jan23> | Total: 102  P Labeled : 71  NP Labeled: 31  (Unseen to the Model) | Precision  65% | Recall  66% |
| Sample set:  <https://tinyurl.com/expr1-reg-jan23> | F1  65% | Accuracy  52% |
| 02 | Exp 2 Jan23 - Regular Dataset - Tokenization 2  6 OM Combined | Notebook:  <https://tinyurl.com/expr2-jan23> | Total: 102  P Labeled : 71  NP Labeled: 31  (Unseen to the Model) | Precision  69% | Recall  63% |
| Sample set:  <https://tinyurl.com/expr1-reg-jan23> | F1  64% | Accuracy  55% |
| 03 | Exp 3 Jan23 - Regular Dataset - Tokenization 3 - 6 OM combined | Notebook:  <https://tinyurl.com/expr3-reg-jan23> | Total: 102  P Labeled : 71  NP Labeled: 31  (Unseen to the Model) | Precision  70% | Recall  78% |
| Sample set:  <https://tinyurl.com/expr1-reg-jan23> | F1  74% | Accuracy  72% |
| 04 | Exp 4 Jan23 - Regular Dataset - Tokenization 4  6 OM combined | Notebook:  <https://tinyurl.com/expr4-reg-jan23> | Total: 102  P Labeled : 71  NP Labeled: 31  (Unseen to the Model) | Precision  0% | Recall  0% |
| Sample set:  <https://tinyurl.com/expr1-reg-jan23> | F1  0% | Accuracy  30% |
| 05 | Exp 5 Jan23 - Regular Dataset - Tokenization 5  6 OM combined | Notebook:  <https://tinyurl.com/expr5-reg-jan23> | Total: 102  P Labeled : 71  NP Labeled: 31  (Unseen to the Model) | Precision  80% | Recall  80% |
| Sample set:  <https://tinyurl.com/expr1-reg-jan23> | F1  78% | Accuracy  72% |
| 7 OM combined dataset experiments | | | | | |
| 06 | Exp 6 Jan23 - Regular Dataset - Tokenization 1  7 OM combined | Notebook:  <https://tinyurl.com/expr6-reg-jan23> | Total: 102  P Labeled : 71  NP Labeled: 31  (Existing data / seen data) | Incomplete Statements | |
| Sample set:  <https://tinyurl.com/7-om-reg-jan23> |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 07 | Exp 7 Jan23 - Regular Dataset - Tokenization 2 –  7 OM combined | Notebook:  <https://tinyurl.com/expr7-reg-jan23> | Total: 102  P Labeled : 71  NP Labeled: 31  (Existing data / seen data) | Precision  76% | Recall  93% |
| Sample set:  <https://tinyurl.com/7-om-reg-jan23> | F1  83% | Accuracy  74% |
| 08 | Exp 8 Jan23 - Regular Dataset - Tokenization 3  7 OM combined | Notebook:  <https://tinyurl.com/expr8-reg-jan23> | Total: 102  P Labeled : 71  NP Labeled: 31  (Existing data / seen data) | Precision  87% | Recall  67% |
| Sample set:  <https://tinyurl.com/expr1-reg-jan23> | F1  76% | Accuracy  71% |
| 09 | Exp 9 Jan23 - Regular Dataset - Tokenization 4  7 OM combined | Notebook:  <https://tinyurl.com/expr9-reg-jan23> | Total: 102  P Labeled : 71  NP Labeled: 31  (Existing data / seen data) | Precision  75% | Recall  71% |
| Sample set:  <https://tinyurl.com/expr1-reg-jan23> | F1  73% | Accuracy  64% |
| 10 | Exp 6 Jan23 - Regular Dataset - Tokenization 5  7 OM combined | Notebook:  https://tinyurl.com/expr10-reg-jan23 | Total: 102  P Labeled : 71  NP Labeled: 31  (Existing data / seen data) | Precision  70% | Recall  87% |
| Sample set:  Sample set:  <https://tinyurl.com/expr1-reg-jan23> | F1  77% | Accuracy  66% |

**RESULTS COMPARISON BETWEEN 6 OM AND 7 OM DATASETS BASED EXPERIMENTS.**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 6 OM Combined  (Unseen data) | 6 OM based results | | 7 OM Combined  (Trained on Existing) | 7 OM based results | | Results comparison | |
| Exp 1 Jan23 - Regular Dataset - Tokenization 1 | Precision  65% | Recall  66% | Exp 6 Jan23 - Regular Dataset - Tokenization 1 | Incomplete Statements | | Incomplete Statements | |
| F1  65% | Accuracy  52% |
| Exp 2 Jan23 - Regular Dataset - Tokenization 2 | Precision  69% | Recall  63% | Exp 7 Jan23 - Regular Dataset - Tokenization 2 | Precision  76% | Recall  93% | Precision  76 – 69 = 7% | Recall  93 – 63 = 30% |
| F1  64% | Accuracy  55% | F1  83% | Accuracy  74% | F1  83 – 64 = 19% | Accuracy  74 – 55 = 19% |
| Exp 3 Jan23 - Regular Dataset - Tokenization 3 | Precision  70% | Recall  78% | Exp 8 Jan23 - Regular Dataset - Tokenization 3 | Precision  87% | Recall  67% | Precision  87 – 70 = 17% | Recall  93 – 63 = 30% |
| F1  74% | Accuracy  72% | F1  76% | Accuracy  71% | F1  76 – 74 = 2% | Accuracy  72 – 71 = 1% |
| Exp 4 Jan23 - Regular Dataset - Tokenization 4 | Precision  0% | Recall  0% | Exp 9 Jan23 - Regular Dataset - Tokenization 4 | Precision  75% | Recall  71% | Precision  75 – 0 = 75% | Recall  71 – 0 = 71% |
| F1  0% | Accuracy  30% | F1  73% | Accuracy  64% | F1  73 – 0 = 73% | Accuracy  64 – 30 = 34% |
| Exp 5 Jan23 - Regular Dataset - Tokenization 5 | Precision  80% | Recall  80% | Exp 10 Jan23 - Regular Dataset - Tokenization 5 | Precision  70% | Recall  87% | Precision  80 – 70 = 10 | Recall  87 – 80 = 7% |
| F1  78% | Accuracy  72% | F1  77% | Accuracy  66% | F1  77 – 78 = 1% | Accuracy  66 – 72 = 6% |