Erratum: “Improving Topic Models with Latent Feature Word Representations”

Dat Quoc Nguyen, Richard Billingsley, Lan Du and Mark Johnson

Abstract

Change in clustering and classification results due to the DMM and LF-DMM bugs.

4.3 Document clustering evaluation

FROM (in the original published article): For example with 40 topics on the TMNtitle dataset, the DMM achieves about 6% higher Purity and NMI scores than LDA.

TO: For example with 80 topics on the TMNtitle dataset, the DMM achieves about 7% higher Purity and NMI scores than LDA.

FROM (in the original published article): on the short text TMN and TMNtitle datasets we obtain 3.6% and 3.0% higher Purity at T = 80.

TO: on the short text TMN and TMNtitle datasets we obtain 6.1% and 2.5% higher Purity at T = 80.

4.4 Document classification evaluation

FROM (in the original published article): In addition, our w2v-DMM model achieves 3.6% and 3.4% higher F1 score than the DMM model on short TMN and TMNtitle datasets with T = 80, respectively.

TO: In addition, our w2v-DMM model achieves 5.4% and 2.9% higher F1 score than the DMM model on short TMN and TMNtitle datasets with T = 80, respectively.

(from a part of Table 10 in the original published article): F1 scores for TMN and TMNtitle datasets.

| Data    | Method  | Method  | λ = 0.6 |
|---------|---------|---------|---------|
|         | T=7     | T=20    | T=40    | T=80    |
| TMN     | DMM     | 0.605±0.023 | 0.724±0.016 | 0.763±0.008 | 0.741±0.005 |
|         | w2v-DMM | 0.619±0.033 | 0.744±0.009 | 0.759±0.005 | 0.777±0.005 |
|         | glove-DMM | 0.624±0.025 | 0.757±0.009 | 0.761±0.005 | 0.774±0.010 |
| Improve |         | 0.019    | 0.003    | 0.023    | 0.036    |
| TMNtitle| DMM     | 0.570±0.022 | 0.650±0.011 | 0.654±0.008 | 0.646±0.008 |
|         | w2v-DMM | 0.562±0.022 | 0.670±0.012 | 0.677±0.006 | 0.680±0.003 |
|         | glove-DMM | 0.592±0.017 | 0.674±0.016 | 0.683±0.006 | 0.679±0.009 |
| Improve |         | 0.022    | 0.024    | 0.029    | 0.034    |

TO: F1 scores for TMN and TMNtitle datasets.

| Data    | Method  | Method  | λ = 0.6 |
|---------|---------|---------|---------|
|         | T=7     | T=20    | T=40    | T=80    |
| TMN     | DMM     | 0.607±0.040 | 0.694±0.026 | 0.712±0.014 | 0.721±0.008 |
|         | w2v-DMM | 0.607±0.019 | 0.736±0.025 | 0.760±0.011 | 0.771±0.005 |
|         | glove-DMM | 0.621±0.042 | 0.759±0.011 | 0.759±0.006 | 0.775±0.006 |
| Improve |         | 0.014    | 0.006    | 0.048    | 0.054    |
| TMNtitle| DMM     | 0.500±0.021 | 0.600±0.015 | 0.630±0.016 | 0.652±0.005 |
|         | w2v-DMM | 0.528±0.028 | 0.663±0.008 | 0.682±0.006 | 0.681±0.006 |
|         | glove-DMM | 0.565±0.027 | 0.688±0.001 | 0.684±0.000 | 0.679±0.004 |
| Improve |         | 0.065    | 0.008    | 0.054    | 0.029    |

(from a part of Table 11 in the original published article): F1 scores for Twitter dataset.

| Data    | Method  | Method  | λ = 0.6 |
|---------|---------|---------|---------|
|         | T=4     | T=20    | T=40    | T=80    |
| Twitter | DMM     | 0.505±0.023 | 0.614±0.012 | 0.634±0.013 | 0.656±0.011 |
|         | w2v-DMM | 0.541±0.035 | 0.636±0.015 | 0.648±0.012 | 0.670±0.010 |
|         | glove-DMM | 0.539±0.024 | 0.638±0.017 | 0.645±0.012 | 0.666±0.009 |
| Improve |         | 0.036    | 0.024    | 0.014    | 0.014    |

TO: F1 scores for Twitter dataset.
FROM (a part of Table 7 in the original published article): Purity and NMI results on the TMN and TMNtitle datasets with the mixture weight $\lambda = 0.6$.

| Data     | Method      | Purity | NMI  |
|----------|-------------|--------|------|
| TMN      | DMM         | 0.632  | 0.445 |
|          | w2v-DMM     | 0.639  | 0.437 |
|          | glove-DMM   | 0.646  | 0.445 |
| Improve  |             | 0.014  | 0.017 |
| TMNtitle | DMM         | 0.598  | 0.353 |
|          | w2v-DMM     | 0.583  | 0.324 |
|          | glove-DMM   | 0.601  | 0.354 |
| Improve  |             | 0.003  | 0.001 |

TO: Purity and NMI results on the TMN and TMNtitle datasets with the mixture weight $\lambda = 0.6$.

| Data     | Method      | Purity | NMI  |
|----------|-------------|--------|------|
| TMN      | DMM         | 0.637  | 0.445 |
|          | w2v-DMM     | 0.623  | 0.426 |
|          | glove-DMM   | 0.641  | 0.440 |
| Improve  |             | 0.004  | 0.019 |
| TMNtitle | DMM         | 0.558  | 0.339 |
|          | w2v-DMM     | 0.552  | 0.329 |
|          | glove-DMM   | 0.586  | 0.343 |
| Improve  |             | 0.028  | 0.005 |

FROM (a part of Table 8 in the original published article): Purity and NMI results on the Twitter dataset with the mixture weight $\lambda = 0.6$.

| Data     | Method      | Purity | NMI  |
|----------|-------------|--------|------|
| Twitter  | DMM         | 0.552  | 0.201 |
|          | w2v-DMM     | 0.581  | 0.232 |
|          | glove-DMM   | 0.580  | 0.230 |
| Improve  |             | 0.029  | 0.038 |

TO: Purity and NMI results on the Twitter dataset with the mixture weight $\lambda = 0.6$.

| Data     | Method      | Purity | NMI  |
|----------|-------------|--------|------|
| Twitter  | DMM         | 0.523  | 0.222 |
|          | w2v-DMM     | 0.589  | 0.243 |
|          | glove-DMM   | 0.583  | 0.250 |
| Improve  |             | 0.066  | 0.028 |