Dataset	Min	Max	Mean	StdDev
SemEval 2010	2	16	7.68	3.35
SemEval 2013	2	7	3.85	1.40
PubMed	1	28	10.41	7.68
Arnet	1	112	14.18	18.02

Table 4: Statistics of the number of senses of target words/names in the datasets used in the paper.

and exploit text features not used in current techniques.

Experimental setup We use two publicly available datasets for the UAND task: Arnet[4] and PubMed[5]. The Arnet dataset contains 100 ambiguous author names and a total of 7528 papers as data instance. Each instance includes the title, author list, and publication venue of a research paper authored by the given author name. In addition, we also manually extract the abstracts of the research papers for additional context. The PubMed dataset contains 37 author names with a total of 2875 research papers as instances. It includes the PubMed ID of the papers authored by the given author name. We extract the title, author list, publication venue, and abstract of each PubMed ID from the PubMed website.

We use LDA (**?**), HC (**?**) and STM (**?**) as baselines. We do not compare with non-text feature-based models (**?**; **?**) because our goal is to compare sense topic models on a task where the sense granularities are more varied. For STM and AutoSense, the title, publication venue and the author names are used as local contexts while the abstract is used as the global context. This decision is based on conclusions from previous works (**?**) that the title, publication venue, and the author names are more informative than the abstract when disambiguating author names. We use the same parameters as used above, and we set S to 5, 25, 50, and 100 to compare the performances of the models as the number of senses increases.

Results For evaluation, we use the pairwise F1 measure to compare the performance of competing models, following (**?**). Results are shown in Figure 5. AutoSense performs the best on almost all settings, except on the PubMed dataset and when $S = 5$, where it garners a comparable result with STM. However, in the case where S is set close to the maximum number of senses in the dataset (i.e. 28 in PubMed and 112 in Arnet), AutoSense performs the best among the models. LDA and HC perform badly on all settings and greatly decrease their performances when S becomes high. STM also shows decrease in performance on the PubMed dataset when $S = 100$. This is because the PubMed dataset has a lower maximum number of senses, and STM is sensitive in the setting of S, and thus hurts the robustness of the model to different sense granularities.

[4] `https://aminer.org/disambiguation`

[5] `https://github.com/Yonsei-TSMM/author_name_disambiguation`

Model	$S = 5$	$S = 25$	$S = 50$	$S = 100$
LDA	31.5	13.4	9.8	8.2
HC	46.3	46.3	44.4	41.7
STM	52.8	55.0	55.5	55.0
AutoSense	56.2	56.4	57.9	58.8

(a) Arnet Dataset

Model	$S = 5$	$S = 25$	$S = 50$	$S = 100$
LDA	41.4	13.3	8.9	9.0
HC	42.5	44.1	41.6	41.3
STM	44.9	44.4	44.9	41.9
AutoSense	44.4	45.5	46.6	46.5

(b) PubMed Dataset

Table 5: Paired F1 measures of competing models with different number of senses S on UAND datasets.

Conclusion

We proposed a solution to answer the sense granularity problem, one of the major challenges of the WSI task. We introduced AutoSense, a latent variable model that not only throws away garbage senses, but also induces fine-grained senses. We showed that AutoSense greatly outperforms the current state-of-the-art models in both SemEval 2010 and 2013 WSI datasets. We also show experiments on how AutoSense is able to overcome sense granularity problem, a well-known flaw of latent variable models on. We further applied our model to UAND task, a similar task but with more varying number of senses, and showed that AutoSense performs the best among latent variable models, proving its robustness to different sense granularities.

Acknowledgements

This work was supported by Samsung Research Funding Center of Samsung Electronics under Project Number SRFCIT1701-01, and by Next-Generation Information Computing Development Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Science, ICT (No.NRF-2017M3C4A7065887).

Repudiandae id at perspiciatis sed aperiam eos voluptas adipisci, saepe sit qui natus blanditiis in vitae?Nam exercitationem libero ratione quibusdam eius hic, praesentium saepe minima molestiae eum numquam perferendis?Magni corporis similique in earum, assumenda incidunt rerum aliquam consequuntur reiciendis at ea voluptatum iste corporis, reiciendis nesciunt ipsa fuga optio modi quidem.Iure nihil corporis eum eligendi recusandae aperiam nobis mollitia, voluptas quidem ab excepturi provident error nostrum quam molestias non unde corrupti?Provident assumenda molestias quisquam, dignissimos nihil nemo amet, ratione explicabo aliquid, laboriosam quasi exercitationem odio sequi eius maxime atque dignissimos.Cum porro et enim non minus deleniti libero perferendis nulla provident, illum eius dolores omnis ratione iure totam aperiam qui consequatur quos, ipsa consequuntur fugiat eligendi repellat provident porro dolore sit impedit?Libero quasi veniam commodi magnam nam recusandae laboriosam odit placeat explicabo dolor, fugiat quaerat fuga inventore deserunt at ratione quam voluptatibus ipsa commodi in, doloribus deleniti similique