Confirming the themes and interpretive unity of Ghazal poetry using topic models

Ehsaneddin Asgari, Marzyeh Ghassemi & Mark Alan Finlayson
Computer Science and Artificial Intelligence Laboratory
Massachusetts Institute of Technology
{asgari, mghassem, markaf}@mit.edu

Abstract

We apply topic modeling to classifying the genre of Ghazal, a form common in Persian poetry. We show that a classifier based on automatically-generated topics exposes important information with only a small performance penalty: the top discriminative topics can be manually aligned with themes prevalent in the associated genres, as identified by scholars of literature. We also weigh in on a long-standing debate about the interpretive unity of Ghazal. In particular, we show evidence that, on the average, Ghazals seem to have interpretive unity at the level of the full poem, as opposed to just at the level of the couplet. Our dataset is a collection of almost 18,000 Ghazal, comprising over 3 million words. The collection contains poems from 30 different poets, and spans nearly 900 years (1080 A.D.–1968).

Generative models tend to perform worse than discriminative models in classification tasks. The hope is, however, that a generative model provides additional insight, by allowing decisions to be understood with reference to the generated explanation. We show that this hope holds true for topic models [1] in two specific cases in the domain of Ghazal, a form common in Persian poetry. First, we show that, for a 3-class genre classification task, the most discriminative topics correspond to known themes prevalent in the associated genres. These themes have been identified by scholars of Persian poetry, and we manually align computationally-extracted topics with these human-identified themes.

Second, we show that classification via topics can provide some insight into actual scholarly debates. Here we focus on *interpretive unity*, a debate among scholars of Ghazal regarding whether the poems are best interpreted at the level of a whole poem, or at the level of a couplet. [2, 3] We train a 30-class topic-model classifier on the Ghazal and use it to do poet attribution (the identification of which poet wrote which poem). As expected, poet attribution is poor when the model is trained using couplets as the document unit, but reasonably good when the model is trained using poems as the document unit. Interestingly, however, when couplets are rearranged (within a poet's poems) to create "scrambled" Ghazals, and the model is trained using these fake Ghazal as the document unit, performance is extremely poor. We take this as evidence that, on the average, Ghazal are coherent, and have interpretive unity, more at the level of the poem than the couplet.

1 Data

The dataset is a collection of Ghazal extracted from the Ganjoor poems website[1]. The collection contains poems from 30 poets, ranging chronologically from Hakim Sanai (1080 A.D.) to Rahi Moayyeri (1968). The collection comprises 17,326 Ghazals containing 156,938 couplets and approximately 3 million words. The dataset is described in detail in [4].

---

[1] http://ganjoor.net/

Ghazals are composed of couplets (which themselves contain two phrases, called hemistichs). Ghazals usually contain around 10 couplets, and couplets contain, on average, about 20 words, for an average of about 200 words per Ghazal.

Ghazals may be classifed based on their genre, which can roughly categorized into four: Khorasani, Iraqi, Indian, and Voghu' (Realist style). [5, Chap. 3] These genres as associated with the 13th-15th centuries, mid-9th–11th centuries, 17th–mid-18th centures, and 16th century, respectively. The labels are merely conventional, and do not necessarily reflect association with a particular region. Ghazals also are rich in a set of traditional metaphors from Persian culture. Some of the most prevalent include: using drinking to refer to experiencing mystical grace; using mirrors or the Biblical figure of Joseph as symbols of beauty; or using a butterfly drawn to a candle to refer to an intensity of love which leads to the loss of one's self. Some scholars of Ghazal hypothesize that couplets can be interpreted independently of the poems in which they are found. Others believe that poems must be interpreted as a whole. We refer to this here as the *interpretive unity* debate.

The dataset was extensively pre-processed before analysis. A previously created lexicon of approximately 1.8 million terms that supports both modern and ancient Persian was used for normalization and lemmatization, as well as several other tools specifically developed for Persian. Pre-processing was done so as to normalize over the change in grammatical structures of the language over the long time period considered (1080 A.D. to 1968). [6] Tokenization was specifically adapted to the structure of the poems, taking advantage of the lexicon and the Ghazals' metre rules to deal with ambiguity inherent in the spacing structures. Details of the lexicon and tool suite may be found in [4].

2 Topics underlying genre classification

Although there are four major genres of Ghazal, one (Khorasani) is extremely rare and so was excluded from the analysis, and therefore this was a 3-way classification problem. First Ghazals were labelled according to the poet's genre, as provided by literary scholars. We removed poets considered to be inter-genre. This resulted in a dataset containing 15 poets, 11,671 Ghazals, 132,260 couplets, and approximately 2.1 million words. Second, we extracted topics from using Latent Dirichlet Allocation (LDA) [1], with poems as the document unit. The LDA model was trained using a bag-of-words assumption for tokens on Mallet v2.0.7. with couplets as the document unit, and then aggregated couplet topic membership across individual Ghazals. The number of topics was chosen based on maximizing model log-likelihood estimated on a validation set using 1x10-fold cross-validation. [7] Each fold was run for 2000 iterations, and hyper-parameter optimization began after 200 iterations. [8] As shown in [4], log-likelihood stabilizes near 500 topics, which is the number we used. For the 3-way classification we used each poem's topic distribution as the feature vector and used Linear-Discriminant-Analysis-based IVEC1 method described in [9], and report metrics based on the 1x10-fold cross-validation averages. The average cross-validation correct rate was 81.69% (Fleiss $\kappa = 0.60$, F_1-score = 0.68, sensitivity = 0.96, specificity = 0.63).

The results of the topic model classification compares favorably with a simple baseline: lemmatized bag-of-word (BOW) features with the IVEC1 classification method and sparse singular value decomposition (SVD) for dimensionality reduction. This baseline obtained a correct rate of 51.82% (Fleiss $\kappa = 0.41$, F_1-score = 0.19, sensitivity = 0.57, specificity = 0.74).

The topic model classification also is proportionally similar to the best genre classification result to date: Argamon et al. [10] reported an 86.6% accuracy when classifying 425 peer-reviewed scientific articles into two different fields, geology or paleontology. They used various linguistic features (relative frequencies of function words, etc.) and a linear kernel SVM for classification.

2.1 Discriminative topics

Table 1 shows the top eight topics which contribute to the discriminant plane for the three genres. Table 2 shows our manual alignment of these topics with common themes and metaphors identified by literary scholars. Note that we have fewer samples from the Voghu' genre than Iraqi and Indian, as most scholars study differences between those two genres.

Table 1: Top eight topics forming the discriminant plane for Iraqi, Indian, and Voghu' genres. Each topic is listed with its most closely associated genre, plus a list of the most probable words.

Dignity (Indian)	**Inattention** (Indian)	**Longing** (Indian)	**Reflections** (Indian)
Effort (همت)	See (نگه)	Footprint (اثر)	Mirror (آینه)
Much (*adj for effort*) (بلند)	Eyelash (مژه)	Lament (ناله)	Flourish (جلوه)
Luck (اقبال)	Inattention (تغافل)	Trackless (بی‌اثر)	Beauty (حسن)
Less (*adj for effort*) (دون)	Watching (تماشا)	To take (بردن)	Bidel (*A poet*) (بیدل)
Dignity (جاه)	Eye, See (نگاه)	To drag (کشاندن)	Face/Meeting (دیدار)
Shame (ننگ)	Saturated[1] (سیر)	Orfi (*A poet*) (عرفی)	Watching (تماشا)
Poverty (فقر)	My God (یارب)	Ambergris[2] (عنبر)	Effacement (محو)
Esteem (عزت)	Intention (منظور)	Ambergris lode (عنبرسارا)	Color (رنگ)

Garden (Iraqi)	**Lover's Dreams** (Iraqi)	**Silence/Speech** (Iraqi)	**Questioning of love** (Voghu')
Meadow (چمن)	Dreaming (خیال)	Grace (فیض)	What? (چیست)
Flower (گل)	Endearment (نوازش)	Speech (حرف)	Who? (کیست)
Nightingale (بلبل)	Impossible (محال)	Discourse (سخن)	Plan (تدبیر)
Rose Arbor (گلستان)	Once in a lifetime (یکبار)	To talk (گفتن)	Fate (تقدیر)
Song birds (مرغ)	Devotee (بنده)	Speech (گفتار)	Love (عشق)
Lament (ناله)	Anguished (مسکین)	Poem (شعر)	Purpose (غرض)
Dawn (سحر)	Lover (عاشق)	Silence (خموشی)	Vahshi (*A poet*) (وحشی)
Garden (باغ)	All night (همه‌شب)	Meaning, Spirit (معنی)	To return (بازگشتن)

[1] Refers to being saturated with a visual experience.
[2] Ambergris is a scent used in perfumes and is often used to stand in for any pleasant scent associated with a desired object.

Table 2: Top discriminative topics manually aligned with analyses by scholars of Ghazal.

Topic	**Genre**	**Analysis**
Dignity	Indian	Some poets emigrated to India to avoid being controlled by the rulers of their native lands; they wanted to gain esteem and dignity through hard work, and not through flattery. [11]
Inattention	Indian	A lover focusing on a beloved woman, but her showing inattention, is a common theme of the Indian genre. [12]
Longing	Indian	Ghazal of the Indian genre often describes the poet searching for the footprint, scent, or emblems of a beloved woman. (In contrast, in the Iraqi genre the beloved doesn't leave any trace of themselves; and in both the Iraqi and Voghu' genres the poet searches specifically just for the face of the beloved woman.) [12]
Reflections	Indian	In the Indian genre mirrors are a key metaphor. [13]
Gardens	Iraqi	In the Iraqi genre, the bird is the "lover" of a flower in a garden. (In contrast, in the Indian genre the flower represents a painting on a carpet.) [12]
Lover's dreams	Iraqi	Women in India were not covered by the Hijab (as in Iran or other Muslim lands) and thus were more "accessible". Dreams were considered inaccessible, and so dreaming is more related to Iraqi genre than Indian genre. [12]
Silence/Speech	Iraqi	Rumi is one the most important Iraqi poets. His nick name is *Khamoush* which means silent. [14]
Questioning of love	Voghu'	In the Voghu' genre, a lover often tries to question the beloved woman and threaten her that he will return from the journey of love. [15]

3 Investigating interpretive unity

There is a debate among scholars of Ghazal whether or not couplets must be interpreted independently or in the context of a whole Ghazal. [2, 3] We designed a task to see whether the topical signature of a poet is encoded in a couplet or in a whole poem. Topic modeling was run as above, except the unit of topic aggregation was varied: topics were aggregated for just couplets, for full Ghazals, and for "scrambled" Ghazals created by interchanging couplets from different poems by the same poet. For the 30-way classification of poem to poet, we used each document unit's topic distribution as the feature vector and used the IVEC1 method as above with 1x10-fold cross-validation. Our comparison was a BOW feature vector on lemmatized data with SVD, and SVM for classification with 1x10 cross-validation. As shown in Table 3, topic modeling shows a distinct advantage: classification accuracy is good for full Ghazals, and much poorer for individual couplets. Importantly, however, the accuracy is worst when scrambled Ghazals are used, which suggests that, on average, a Ghazal itself has interdependent couplets and should be seen as a whole unit.

Table 3: Poet attribution accuracies using aggregation over different document units.

Unit	Topic Model (%)	BOW-SVM (%)
Ghazals	**55.0**	23.1
Couplets	23.9	23.0
Scrambled	2.8	5.5

4 Contributions

We have made two contributions in this paper. First, we show that the topics most important for classifying Ghazal into genres correspond to actual metaphors and symbolisms identified by literary scholars. This shows that topic models, while performing nearly as well as a state of the art genre classfier, give us additional information that is useful for more detailed and in-depth understanding of genre. Second, we weigh in on the interpretive unity debate by showing that poet attribution accuracy is severly penalized (indeed, almost eliminated) when the classifer is trained on topics extracted from a poet's scrambled poems. We take this as evidence that Ghazal's should be, on the average, interpreted at the level of the whole poem, rather than just individual couplets.

References

[1] D. M. Blei. Probabilitic topic models. *Communications of the ACM*, 55(4):77–84, 2012.

[2] T. Malmir. Cohesive structure of Hafez's Ghazals. *Literary Techniques*, (1):41–56, 2009.

[3] M. Heydari. Studying the structure of Persian Ghazals. *Language and Linguistics*, (31):24–38, 2007.

[4] E. Asgari and J.-C. Chappelier. Linguistic resources & topic models for the analysis of Persian poems. In *Proceedings of the Second Workshop on Computational Linguistics for Literature (CLfL 2013)*, pages 23–31, Atlanta, Georgia, 2013.

[5] A. Tamīmdārī. *The Book of Iran: A History of Persian Literature: Schools, Periods, Styles, and Literary Genres*. Alhoda, Tehran, 2002. Translated by Ismā'īl Salāmī.

[6] P. N. Xañlari and E. Mostasharniya. *A Historical Grammar of Persian Language*. Tose'eye, Tehran, 2009.

[7] W. Buntine. Estimating likelihoods for topic models. In *Proceedings of the First Asian Conference on Machine Learning (ACML'09)*, pages 51–64, Nanjing, China, 2009.

[8] H. Wallach, D. Mimno, and A. McCallum. Rethinking LDA: Why priors matter. In *Advances in Neural Information Processing Systems 22 (NIPS'09)*, pages 1973–1981, 2009.

[9] E. Singer, P. Torres-Carrasquillo, D. Reynolds, A. McCree, F. Richardson, N. Dehak, and D. Sturim. The MITLL NIST LRE 2011 language recognition system. In *Proceedings of Speaker Odyssey 2012*, pages 209–215, Singapore, 2012.

[10] S. Argamon, C. Whitelaw, P. Chase, S. R. Hota, N. Garg, and S. Levitan. Stylistic text classification using functional lexical features. *Journal of the American Society for Information Science and Technology*, 58(6):802–822, 2007.

[11] S. Ghahreman. Indian genre and the symbol of opposition. *Literature and Linguistics*, (41):65–82, 2009. In Persian.

[12] A. Zadeh. Innovations in Bidel's poems. *Literature and Liguistics*, (10):139–165, 2008. In Persian.

[13] M. R. Shafiee Kadkani. *Poet of mirrors*. Agaah, Tehran, 2007.

[14] B. Fakhroleslam. Why did Rumi has choose 'silent' as his nickname? *Cultural Keyhan*, (176):20–21, 2001. In Persian.

[15] T. RazmJoo. A survey of the Voghu genre in Persian literature. *Monthly Book of Literature*, 162:80–84, 2011. In Persian.