

Using Wine Descriptions to Predict Varietal

(and the other way around)

Clara Meister, Tim Aiken

MOTIVATION

Parsing wine reviews is an interesting subset of NLP in that the reviews are 1) mostly positive and 2) highly subjective. We chose to undertake this challenge as a way of experimenting with this interesting subset of language and seeing if we could derive order from the (flowery, tannic) chaos.

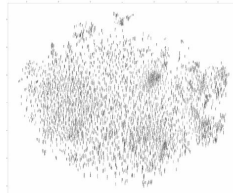
PROBLEM DEFINITION

We take a database of 150,000 wine reviews and attempt to predict the varietal (type of wine grape) being described based on the text of the reviews. We also attempt to predict the wine review (English words) based on the varietal and rating (wines are scored on a scale of 80-100).

APPROACH

- **Word embeddings:** We use GloVe embeddings to encode our text. The GloVe matrix was trained for this specific problem on a co-occurrence matrix generated from the training data.
- **Many-to-One RNN:** A single layer Bidirectional LSTM is used to predict varietal. This is a classification problem that takes an input of text encodings over time and produces a softmax output with 31 possible categories (30 varietals and 'other').
- **One-to-Many RNN:** A single layer LSTM is used to predict review words from varietal and point score. This is a classification problem that initializes the LSTM activation with varietal and point score, then at every step it takes the previous word in the review and predicts the next one.

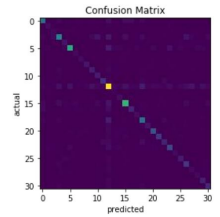
GloVe



t-SNE visualization

Varietal Prediction

- Overall 65% accuracy on test set.
- Classifies 82% of Chardonnay correctly.
- Predicts 78% of Tempranillo incorrectly.



Review Prediction

Input: Pinot Noir, 80 / 100 pts (bad score)
Output: Very sweet, almost syrupy. A pinot that's the acidity accentuates and it's soft which doesn't get there. It's that will the raspberry cream note that taste overwhelm, it this for those tannins for aging yet this rich too much like what it has more power in their bottles sh, and are best at and try holding down.

Analysis

Both our varietal prediction and review prediction models perform much better than expected. We estimate a rough human baseline for varietal prediction to be 10%, so we vastly exceed that level of performance. The review prediction model learned interesting details of varietal, quality (point score), and review language, qualitatively producing very compelling results.