

TAITE SANDEFER

DATA SCIENTIST



512-965-7295



taite.sandefer@gmail.com



[/in/tsandefer](https://in.linkedin.com/in/tsandefer)



[/tsandefer](https://github.com/tsandefer)



Denver, CO

PROJECT SPOTLIGHT

NEURAL MACHINE TRANSLATION: EXPLAINING THE MEANING BEHIND LYRICS

Used lyric and annotation data from Genius.com API to train a character-based Seq2Seq LSTM model on AWS EC2 instance with GPU. Deployed Flask app on Heroku with trained model that generates new explanations for user-input lyrics. [GitHub](#).

Tools: Seq2Seq with RNN/LSTMs, Keras, Tensorflow, AWS EC2, Flask, CSS, HTML, Heroku

DOC2VEC AND ANNOTATED LYRICS: ARE THEY "GENIUS"?

Gathered lyric and annotation data from Genius.com API to explore whether the cosine similarity of their Doc2Vec vector representations could be useful in developing an automated system to accept/reject newly proposed annotations. Hypothesis test across true and falsely matched pairs yielded a significant p-value of $1.1e-171$. [GitHub](#).

Tools: Gensim's Doc2Vec, t-SNE, MongoDB, SciPy's two-tailed hypothesis test

DOES WORK-LIFE BALANCE MATTER?

Leveraged student outcome and participation data across two subjects to predict social behaviors using logistic regression modeling with cross validation. Final models yielded F1 scores of 0.461 and 0.501, with an AUC of 0.632 and 0.521, respectively. [GitHub](#).

Tools: Scikit-Learn's Logistic Regression, k-Fold Cross Validation, GridSearch

SELECT EXPERIENCE

PROJECT MANAGEMENT INTERN

ERCOT // Taylor, TX // MAY 2017 – DEC 2017

- Analyzed data gathered from conducting in-depth internal research using Qualtrics, Excel, and Tableau software to inform and support recommendations for change in overhaul of Project Lessons Learned procedures
- Reduced time spent on Project Closing Procedures by 90%
- Presented all proposed plans to executive leadership for decisions and approvals

MARKETING INTERN

Bare Bones Marketing // Austin, TX // SEP 2016 – MAY 2017

- Wrote marketing proposals for potential new clients, with an 88% conversion rate
- Created content and strategy for over 200 social media posts across 3 campaigns
- Implemented CRM system that automated all regular client communication

INSTRUCTOR, OFFICE MANAGEMENT

Executive Watersports, Ltd. // Austin, TX // JUN 2012 – AUG 2016

- Taught over 100 campers (ages 5-14) how to waterski and/or wakeboard
- Managed regulatory compliance and communication with over 250 clients
- Conducted daily maintenance routines for all boats to ensure peak performance

KEY SKILLS

PYTHON

NumPy, Pandas, SciPy, Scikit-Learn, Keras, TensorFlow, PyTorch

DATA PIPELINE & ANALYSIS

Collection & Cleaning, EDA, Feature Engineering, A/B Testing

Visualization: Matplotlib, Tableau, Seaborn, t-SNE

NATURAL LANGUAGE PROCESSING

*Libraries: NLTK, spaCy, Gensim, t-SNE
Representation: Tf-idf/BoW, Word2Vec, Doc2Vec, Seq2Seq*

Topic Modeling: NMF, LDA

Generative Modeling: RNN/LSTM

SOTA: BERT, Transformer (Relative Self-Attention Mechanisms)

MACHINE LEARNING METHODS

Supervised: Linear, Logistic, LASSO, Ridge, & Elastic Net Regression; k-NN, Random Forest, Boosting

Unsupervised: K-Means, Hierarchical Clustering, PCA, SVD, NMF, LDA

DATABASE, CLOUD, & PRODUCT

PostgreSQL, MongoDB, Spark, AWS, Flask, Heroku, CSS, HTML, Unit Tests

PROJECT MANAGEMENT & COLLABORATIVE CODING

Git, GitHub, Agile Workflow, JIRA

EDUCATION

DATA SCIENCE IMMERSIVE

Galvanize // Denver, CO
JAN 2019 – APR 2019

B.A. IN ECONOMICS,

Minor in Rhetoric & Persuasive Writing
University of Texas // Austin, TX
AUG 2016 – AUG 2018