Abstract

# [Bag of Words Meets Bags of Popcorn](https://www.kaggle.com/c/word2vec-nlp-tutorial) by Kaggle

# Nicolò Brandizzi

# Supervised, classificazione

# [Dataset](https://www.kaggle.com/c/word2vec-nlp-tutorial)

# Vettore delle x originariamente composto da Stringhe poi vettorializzato, binario

# Vettore delle Y binario

# Valutazione di una recensione di un film che può essere positiva o negativa

# Random Forest, SVM, chi2

# ??

# Media dell’errore nelle predizioni

# Nessuna validazione, il dataset è dato gia diviso in test e train

# Random Forest =85%, SVM=89%

# Python

# [Sklearn](http://scikit-learn.org/stable/index.html): per i classificatori, la vettorializzazione delle stringhe e la dimensionality reduction.

# [Pandas](http://pandas.pydata.org/pandas-docs/stable/index.html): per la creazione e manipolazione dei datasets.

# [MatplotLib](http://matplotlib.org/): per la creazione dei grafici