

# PRODUCTION VALUE

## Identifying Record Producers from Audio Data

Maxwell Fisch

### Background

A record producer controls the creation of a music album. Each music producer leaves a sonic fingerprint on every album they produce. Production Value seeks to identify and quantify that signature. This model could be used for:

- **Music discovery:** aiding Spotify and Pandora users in finding music they like.
- **Music publishing:** helping record labels identify and distribute royalties to song collaborators.

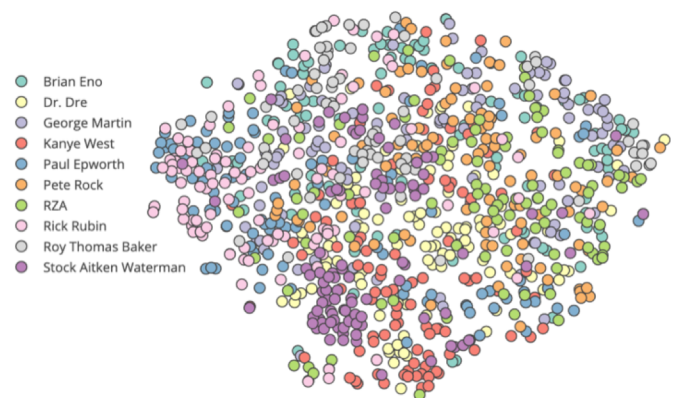
### Approach

- 1000 songs from 10 producers (100 songs each) were chosen.
- 30-second snippets of audio from Spotify's API were processed and transformed into Mel-Cepstral Frequency Coefficients (MFCCs) – a featurization that roughly translates to timbre.
- Principal Component Analysis (PCA) reduced dimensionality (24k features → 12).
- K-Nearest Neighbors (KNN) classification was used to classify the record producers.

### Results and Analysis

- KNN model multiclass accuracy (10 balanced classes) was 44% compared to a baseline of 10%.
- Producers and songs cluster in MFCC-space. Some producers have characteristic sound (e.g. Stock Aitken Waterman), others are diverse (e.g. George Martin).
- Future improvements:
  - Deconvolution of variables (e.g. Artist, Album, Genre) via nuanced feature engineering.
  - Neural Network Model with Increased Scale (>100k songs)

2D t-SNE Plot of Producers



Confusion Matrix

	Brian Eno	Dr. Dre	George Martin	Kanye West	Paul Epworth	Pete Rock	RZA	Rick Rubin	Roy Thomas Baker	Stock Aitken Waterman
Brian Eno	13	2	7	2	4	1	2	0	0	2
Dr. Dre	0	11	4	3	1	3	2	0	1	6
George Martin	2	1	8	0	1	2	4	0	1	3
Kanye West	0	4	2	14	1	0	0	0	0	4
Paul Epworth	0	0	4	1	16	0	0	2	1	4
Pete Rock	3	4	5	0	0	12	6	0	1	0
RZA	4	5	1	1	1	3	9	0	1	3
Rick Rubin	4	2	2	0	11	2	0	12	1	3
Roy Thomas Baker	1	0	6	1	1	1	1	0	8	2
Stock Aitken Waterman	0	1	1	1	1	1	1	0	5	24

### Tech Stack:



linkedin.com/in/maxwellhfisch |



github.com/mhfisch |



@Maxwell\_Fisch



maxwellfisch@gmail.com | (805) 982-0955

Project Links: <http://github.com/mhfisch/production-value> | <http://www.production-value.com>

# MAXWELL FISCH

(805) 982-0955 | [maxwelfisch@gmail.com](mailto:maxwelfisch@gmail.com) | [linkedin.com/in/maxwelfisch](https://www.linkedin.com/in/maxwelfisch) | [github.com/mhfisch](https://github.com/mhfisch)

## DATA SCIENTIST

Data Scientist with an extensive background as a Root Cause Failure Analyst with a history of highly analytical work. Proven success in combining modern data science tools with a wide breadth of engineering skills. Demonstrated career history of exceeding project deliverables and corporate performance metrics.

Python Pandas Numpy | SQL | MongoDB | AWS Spark | Matplotlib Plotly Seaborn | SKlearn SciPy

## DATA SCIENCE EXPERIENCE AND PROJECTS

### Production Value

2019

#### **Data Science Capstone Project**

Identifying Record Producers from Audio Data Using Machine Learning

- Accurately predicted record producer 4 times better than the baseline.
- Built a platform that queries audio data from Spotify and returns likely producers and similar songs based on a KNN model. See the whole project at [github.com/mhfisch/production-value](https://github.com/mhfisch/production-value).

### Galvanize

2019

#### **Data Science**

12 weeks of immersive data science coursework, case studies, and projects.

- Mastered fundamentals of modeling, machine learning, hypothesis testing, and linear algebra.
- Extensive experience with Python libraries, including: Numpy, Pandas, pyspark, PyMongo, Scikit-learn, Statsmodels, Scipy, Matplotlib, Seaborn, Nltk (Natural Language Toolkit), and BeautifulSoup.
- Learned industry workflow tools and best practices, utilizing Git, Unix, Anaconda, bash scripting, SQL, MongoDB, AWS, and Spark.

## PROFESSIONAL WORK

### Western Digital

2016 – 2018

#### **Senior Engineer**

Determined root cause in failure analysis for customer return hard drives via materials science and data analysis.

- Published engineering reports and communicated failures to lab managers, customers, and executives.
- Isolated failure sources on flagship product by utilizing materials characterization resulting in 75% reduction in failure.

### Levi Lab, UC, Santa Barbara

2012 – 2015

#### **Graduate Research Fellow**

Built apparatus and designed experiments for measurement of high-temperature water vapor transport of Ca-Mg-Al-Si (CMAS) Oxides.

- Demonstrated a novel vapor-transport CMAS ingress mechanism using a custom experimental rig.
- Analyzed infiltration of CMAS oxides with thermal barrier coatings via materials characterization and microscopy.

### Cooper Bussmann

2011 – 2011

#### **Engineering Intern**

Conducted a design of experiments for electrolytic super capacitor materials and suggested design improvements.

- Optimized super capacitor capacity and reduced leakage through a DOE varying materials selection and capacitor design.

### Maboudian Lab, UC Berkeley

2010 – 2012

#### **Undergraduate Researcher**

Qualified failure mechanisms of a MEMS devices

- Presented results at a conference
- Co-authored three papers.

## EDUCATION

**Master of Science in Materials**, UC Santa Barbara

NSF Graduate Research Fellow. Overall GPA: 3.79/4.00

**Bachelor of Science in Chemical Engineering and Materials Science**, UC Berkeley

Certificate in Technology and Entrepreneurship from Center for Entrepreneurship and Technology.

Regents' and Chancellor's Scholar, Overall GPA: 3.70/4.00

**Extra Interests:** Theater, Music, Cooking, Board Games