title*
subtitle

Michaela Drouillard

11/1/23

1 Introduction

Shortly after the release of Taylor Swift's *Midnights* album, a video of the Youtuber Caleb Gammen, in which he guesses which songs on Midnights were produced by Antonoff within seconds (@calebgamman 2022), went viral. The tweet reads: "im able to instantly detect if jack antonoff worked on a song due to a visceral hatred of his production style". He nails it. Only one song on the album wasn't produced by Antonoff, and he guesses it.

Antonoff began his career as a guitarist in fun., and is the frontman in Bleachers. He has produced and co-produced music for Lorde, Taylor Swift, Florence and the Machine, The Chicks, Clairo, The 1975, Grimes, Zayn, Pink, Lana Del Rey, Olivia Rodrigo, the Minions: The Rise of Gru Soundtrack, and more. Antonoff is prolific, to the extent that it is starting to bother people. As Andrew Marantz writes in the New Yorker: "When his band releases an album, the world responds politely. When he produces one by Lorde or Lana Del Rey or Taylor Swift, the world wobbles on its axis" (Marantz 2022).Can Spotify's audio features dataset help us determine what makes Antonoff music so Antonoff-y?

I spoke to Gamman to better undertstand makes a song Antonoff-ish, and to try to understand his reaction to it. In trying to understand how to describe Antonoff's sound, I used the Spotify API's audio features dataset to explore difference machine listening metrics, and build a model to recreate a digital version of Gamman's classification excercise. I also spoke with Glen McDonald, an engineer at Spotify, from whom I learned that the `valence` and `danceability` variables in the Spotify API were created by hiring college interns and having them tag whether songs were positive or danceable.

I used two models to investigate the data and ask the question: What is Antonoff's sound? Can we pin down style using the quantitative metrics offered by Spotify?

---

*Code and data are available at: https://github.com/michaeladrouillard/spotify_pop.

1

The first model, a logistic regression, draws on the dataset of artists who've collaborated with Antonoff, with an extra binary variable to tag whether or not the song was an Antonoff production. With this model, we can get a first glimpse at variables that are stronger predictors of Antonoff-iness, along with a better understanding of whether the Spotify data can trace Antonoff's influence on an artist's sound over time.

However, logistic regression assumes a linear relationship between the dependent and predictor variables, and is sensitive to multi-collinearity. In music data, we would expect for certain variables to be correlated (for instance, loudness and energy). We also cannot assume that an artists' style can be measured in the linear growth of these metrics. Although the logistic regression has a strong accuracy, it may be detecting a change in genre, rather than a change in style, across artists whose sound turns more and more pop while collaborating with Antonoff.

I built a second model, a Random Forest, to complement the insights drawn from the logistic regression. The Random Forest trains on the full data set, including other major pop producers. Random Forests don't assume linear relationships, and aren't sensitive to multicollinearity. In fact, using the tk package in R, we can visualize and analyse the relationships between different variables. The variance in correlation scores across producers can be taken as an approximation for style, since this variance helps the model distinguish between producers.

In the discussion section, I listen to the songs with the highest danceability scores, and make an argument that this `danceability` feature, which has so much uncertainty cooked into it, is really capturing something like visceral-ness. Against the odds, we can capture something about Antonoff's music with these variables, and it might, in the end, just be that subjective "ick" that some people feel, and others enjoy, and that we're all just trying to describe in using different methods and tools.

2 tk: Lit Review

3 Data

The audio feature data was pulled using Spotify API, which I accessed using the `spotifyr` package in R (Thompson et al. 2022). I used the `get_artist_audio_features` function to acquire data on Lorde, Taylor Swift, St. Vincent, Lana Del Rey, The Chicks, Florence and the Machine, and Bleachers discographies. I threw in HAIM, Marina and the Diamonds, Maggie Rogers, Sharon Von Etten, and Mitski in too, since they've never collaborated with Jack Antonoff but are more or less part of the same pop sphere.

I also scraped the discographies of comparable contemporary pop producers' Wikipedias. The producers include Epworth, Rechstaid, Max Martin, Nowels, Joel Little, and Greg Kustin. (cite Carl here?) These producers were chosen based on their popularity, and also because some of them collaborated with artists who also collaborate with Antonoff.

The data was manually validated, and live performances, karaoke editions, international versions or translations, and remix albums were removed. Where deluxe albums were available, original albums were deleted to avoid duplicate rows.

3.0.1 Artist Selection

Of the artists who did collaborate with Antonoff, I chose these artists because they've produced multiple albums, and at least one album with Antonoff, as opposed to just individual songs.

3.0.2 Data Cleaning

I joined these data sets using the `rbind` function in base R, and created an additional variable called "jack", which contains a 1 if Jack Antonoff produced or co-produced the song, and a 0 if he had nothing to do with it. I got the information on which songs Antonoff did or did not produce from the "Jack Antonoff Production Discography" Wikipedia page (Wikipedia 2023).

3.0.3 Audio Features Dataset

The `get_artist_audio_features` provides a profile of each song, according to variables developed by Spotify, which makes for simple comparisons across artists. The final audio features in this study data set contains the variables: `artist_name`, `track_name`, `energy`, `danceability`, `key`, `loudness`, `mode`, `speechiness`, `acousticness`, `instrumentalness`, `liveness`, `valence`, `tempo`, `jack`.

From the `get_artist_audio_features` documentation ("Get Track's Audio Features" (n.d.)):

- **acousticness**: A confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic.

- **danceability**: Danceability describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable.

- **energy**: Energy is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy. For example, death metal has high energy, while a Bach prelude scores low on the scale. Perceptual features contributing to this attribute include dynamic range, perceived loudness, timbre, onset rate, and general entropy.

- **instrumentalness**: Predicts whether a track contains no vocals. "Ooh" and "aah" sounds are treated as instrumental in this context. Rap or spoken word tracks are clearly "vocal". The closer the instrumentalness value is to 1.0, the greater likelihood the track contains no vocal content. Values above 0.5 are intended to represent instrumental tracks, but confidence is higher as the value approaches 1.0.

- **speechiness:** Speechiness detects the presence of spoken words in a track. The more exclusively speech-like the recording (e.g. talk show, audio book, poetry), the closer to 1.0 the attribute value. Values above 0.66 describe tracks that are probably made entirely of spoken words. Values between 0.33 and 0.66 describe tracks that may contain both music and speech, either in sections or layered, including such cases as rap music. Values below 0.33 most likely represent music and other non-speech-like tracks.

- **key**: The key the track is in. Integers map to pitches using standard Pitch Class notation. E.g. 0 = C, 1 = C /D , 2 = D, and so on. If no key was detected, the value is -1.

- **liveness**: Detects the presence of an audience in the recording. Higher liveness values represent an increased probability that the track was performed live. A value above 0.8 provides strong likelihood that the track is live.

- **loudness:** The overall loudness of a track in decibels (dB). Loudness values are averaged across the entire track and are useful for comparing relative loudness of tracks. Loudness is the quality of a sound that is the primary psychological correlate of physical strength (amplitude). Values typically range between -60 and 0 db.

- **mode:** Mode indicates the modality (major or minor) of a track, the type of scale from which its melodic content is derived. Major is represented by 1 and minor is 0.

- **tempo:** The overall estimated tempo of a track in beats per minute (BPM). In musical terminology, tempo is the speed or pace of a given piece and derives directly from the average beat duration.

- **valence:** A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry).

It's hard to gauge exactly what the variables `danceability` or `valence` mean, even using the documentation. For instance, `valence` is described as "A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad,

depressed, angry)". `Danceability` is described as "how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable". Neither of these definitions fully cover what metrics and thresholds go into something like `danceability`.

3.1 Interview with Glen McDonald

I interviewed Glen McDonald, a Principal Engineer at Spotify who worked at The Echo Nest, a music intelligence start-up that was acquired by Spotify in 2014. Macdonald told me that `valence` and `danceability` were created by giving thousands of examples to college interns and asking them to tag whether a song was positive or gloomy, or danceable or un-danceable. Variables like energy or instrumentalness, on the other hand, were determined primarily through machine listening techniques. The human subjectivity part only came in when it came to tweaking the training data[1]:

"You could imagine writing a formula for energy that combines like loudness and tempo and degree of harmonic variation or something. So it was [machine learning], but that one's more like a human helping a machine figure out a formula. Whereas valence is teaching the machine to try to reproduce a purely human thing. Same with danceability. I mean, danceability is whether a human can dance to it. So like, the machine's not gonna dance, so the machine can't have any opinion on that. The computer could have an opinion on energy. And the computer can definitely have an opinion on loudness. So there's a spectrum from, you know, loudness as purely analytical, and then energy is a little like loudness, with a little more subjectivity. And then danceability and valence are purely subjective."

Glen confirmed that this process didn't know anything about lyrics, which makes valence a particularly difficult variable to build. It could, in theory, pick up on aspects of vocal performance, but it doesn't know anything about languages or words or the meanings of songs. Take an upbeat, happy-sounding Elliot Smith song with devastating lyrics: the machine might register it as happy, where human listeners understand that it's sad. Add to that that two humans might even disagree on the song's valence.

"Plus, we have the confounding factor of like, a song that seems happy today could be sad tomorrow because the singer was killed in a plane crash. The song didn't change, but our world changed and our reactions changed."

Macdonald suggests combining energy and valence to create quadrant, which usually works "fairly well". High energy and high valence could be generally happy, cheerful, upbeat. Low energy, low valence could be sad or downbeat. High energy, low valence is sort of angry. Low energy, high valence is serene or calming.

---

[1]My favourite example Glen gave was that the ML kept giving bluegrass songs insanely high `speechiness` ratings. Because there were no banjos in the training data, they were registered as human speech. (This is hilarious.) They had to go back and add more songs with banjos to the training data.

In **?@fig-quadvalence**, it looks like songs produced by Antonoff are a little more clustered towards the lower end of valence, and across the spectrum of energy, although the variance is large.

In **?@fig-quaddanceability2**, it looks like Antonoff-produced songs err towards higher danceability, across the spectrum of energy.

These quadrants are a start, but they don't suggest any defining Antonoff-iness characteristic underlying each of those blue points, which are scattered pretty broadly.

By comparing cumulative distribution functions of each variable based on whether a song was produced by Antonoff, we can see that some of these variables do capture some sort of underlying Antonoff-iness (**?@fig-cdf**). For instance, while liveness, instrumentalness, and speechiness are relatively similar, we can see clear differences in danceability, energy, loudness, acousticness, and valence.

Another way to gauge the Antonoff-ification of pop music could be to look at how artists' sounds have changed over time, and whether there is any correlation with Antonoff. These plots are here to get a glimpse at pop's evolution over time, and where Antonoff fits in that picture. Each point represents a song.

In **?@fig-energy**, it looks like there's a downward trend in energy scores in artists who work with Antonoff, but that these downward trends are also somewhat in line with the artists' energy trajectories anyway, which could be indicative of broader trends in pop. Specifically, Lana Del Rey and Taylor Swift. They have enough albums recorded prior to collaborations with Antonoff to be able to observe this. Sharon Von Etten, Maggie Rogers, and Mitski, who have never worked with Antonoff, all have upward trends.

There aren't any significant correlations in danceability (**?@fig-danceability**) .

In **?@fig-valence**, it looks like Lorde has the only significant change in valence scores, which tracks with my understanding her albums. *Melodrama* had piano and break ups, and *Solar Power* had a music video of Lorde prancing on the beach in a yellow bikini with Phoebe Bridgers and Clairo.

Again in **?@fig-acousticness**, Lorde's acoustic Solar Power album jumps out. Taylor Swift, Lorde and HAIM also have significant-seeming upward trends in `acousticness`, while Mitski, Sharon Von Etten and Maggie Rogers see decreasing `acousticness`.

In **?@fig-instrumentalness**, St. Vincent is the only plot with any significant trend in `instrumentalness`.

Finally, in **?@fig-tempo**, there aren't many significant trends in tempo, except for Antonoff's collaborations with Lana Del Rey and Taylor Swift.

We've observed subtle distinctions in variables associated with songs produced by Jack Antonoff, as evidenced by our cumulative distribution functions, quadrant analyses, and minor trends in the evolution of artists' sound characteristics, particularly energy. Employing a

logistic regression will then: a) determine the likelihood of a song being produced by Antonoff based on these identified variables, and b) discern which variables are most influential in shaping this classification.

4 Logistic Regression

4.1 Data Preparation

The binary response variable, called `is_antonoff`, has a value of 1 if the song was produced by Antonoff, and a 0 otherwise. There are 10 predictor variables, all of which are continuous: **danceability**, **energy**, **loudness**, **mode**, **speechiness**, **acousticness**, **instrumentalness**, **liveness**, **valence**, and **tempo**.

The formula for this logistic regression can be written as:

$$\Pr(y_i = 1) = \text{logit}^{-1}\left(\beta_0 + \beta_1 x_i\right)$$

where $\Pr(y_i = 1)$ is the probability of a song being produced by Antonoff, β_0 is the intercept, and $\beta_1, \ldots, \beta_{10}$ are the coefficients corresponding to the predictor variables.

Figure 1 gives us an idea of the mean value of each of these features for Antonofff and non-Antonoff-produced tracks.

After some minor pre-processing using the `tidyverse` R package (Wickham et al. 2019). I followed Silge (2023) to build a model that deals with the imbalanced classes.

I used the `tidymodels` to fit the model (Kuhn and Wickham 2020). To deal with the imbalanced classes, I downsampled the majority class using the **step_downsample()** function from the **themis** package. Then I used 5-fold cross-validation with stratification by the **jack** variable to estimate the performance of the model.

The `workflow_set()` function lets us hold different sets of feature engineering together and use tuning to evaluate them all at once. I need to evaluate two elements: a basic model, and the model that was downsampled using the **themis** package.

I used **rank_results()** from the **yardstick** package to rank the models based on accuracy, log loss, sensitivity, and specificity. After testing both basic and downsampled versions of the model, the basic version yielded much higher accuracy score (0.918 basic versus 0.743 downsampled), although a lower roc auc score (0.746 basic versus 0.821 downsampled).

The version with the original data remains consistently high in accuracy, even with increased regularization. It starts low with in mean log loss and sensitivity, but increases with increased regularization. Compared to the downsampled data, which starts high but drops significantly in accuracy and sensitivity, I chose to use the original data.
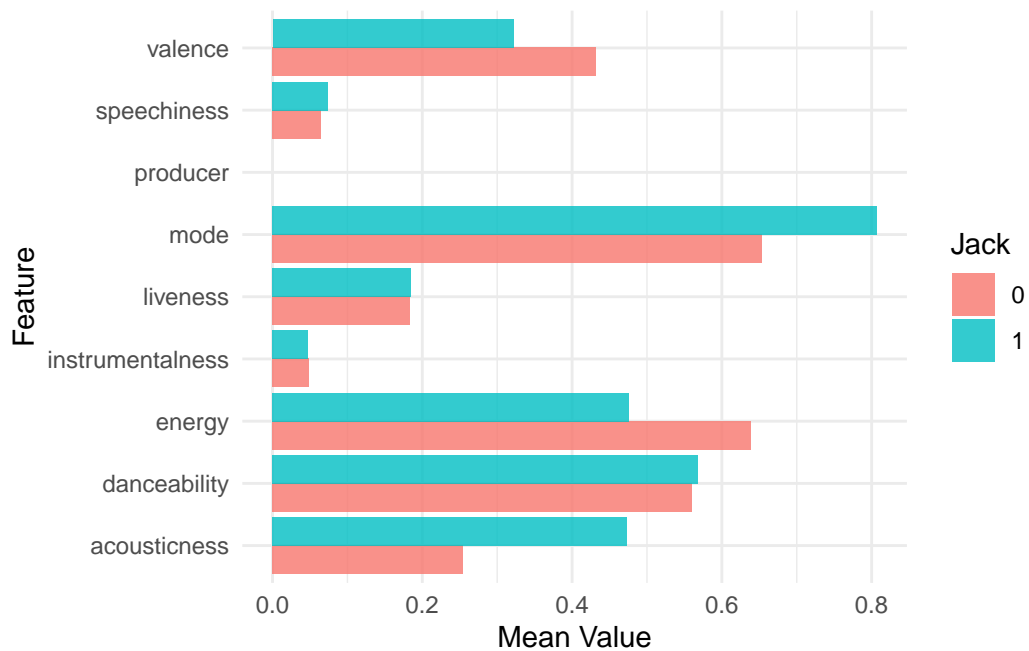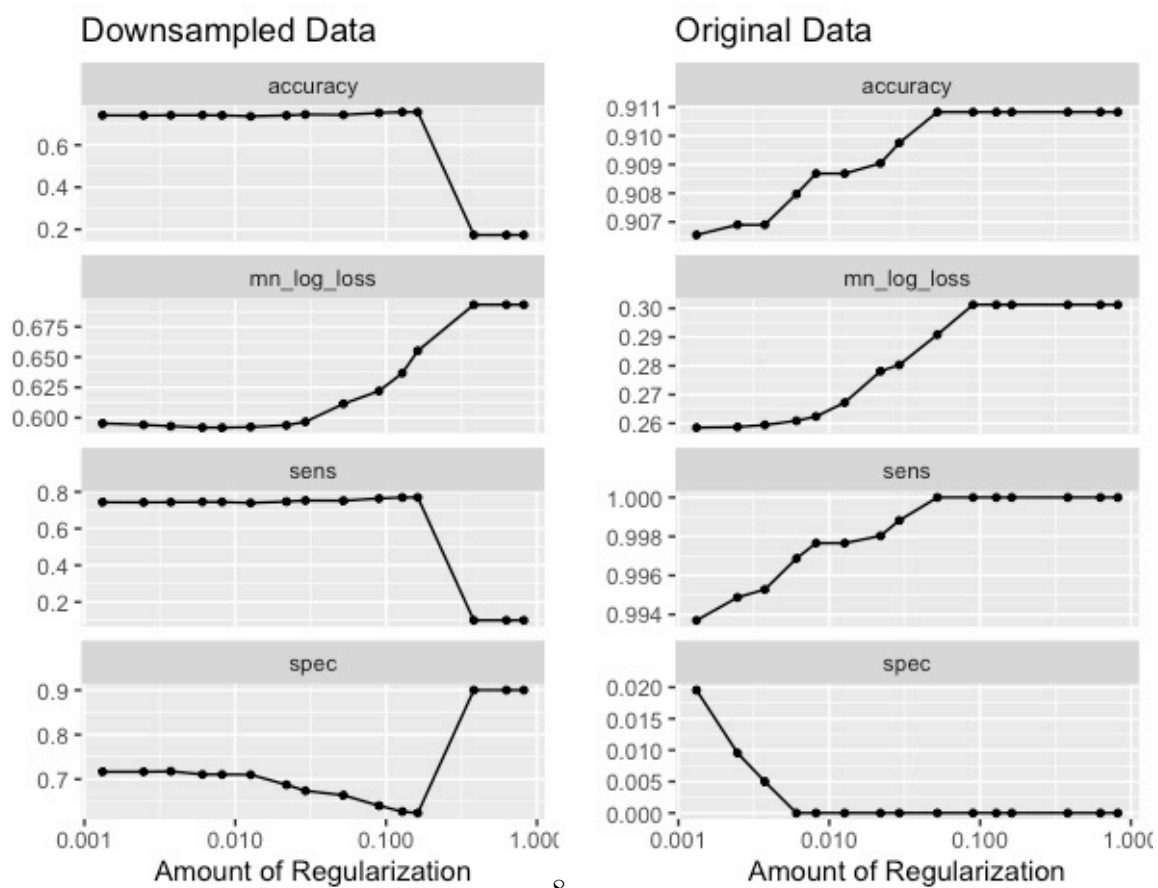
Figure 1: Comparing the Mean Values of Each Variable

4.2 Results

The model's accuracy score is 0.918, and the ROC AUC score is 0.746.

Table 1: Confusion Matrix for Logistic Regression Results

	Actually Positive (1)	Actually Negative (0)
Predicted Positive (1)	848	76
Predicted Negative (0)	1	0

Table 2 shows the different variables' importance on determining whether a song was had any Antonoff involvement.

Table 2: Importance Scores of Each Variable in Determining Antonoff's Tracks

Variable	Importance	Sign
danceability	4.1421839	POS
valence	2.5591576	NEG
instrumentalness	2.3092982	NEG
acousticness	1.4436158	POS
liveness	0.7892735	POS
mode	0.7759733	POS
energy	0.6453724	POS
speechiness	0.3684891	POS
loudness	0.1785005	NEG
tempo	0.0041666	POS

`Danceability`, with a positive importance score of 4.142, had the highest level of importance in determining whether a song was produced or co-produced by Jack Antonoff. `Valence` comes in second at 2.559, with negative score. The third, `instrumentalness` correlates negatively with Antonoff tracks with a a negative score of 2.309.

Table 3: Top 20 Most Danceable Antonoff-Produced Songs

Artist Name	Danceability	Track
Taylor Swift	0.897	I Think He Knows
Taylor Swift	0.875	Vigilante Shit
Florence + The Machine	0.852	Heaven Is Here
Bleachers	0.838	I Wanna Get Better - Demo Version
Taylor Swift	0.824	Cornelia Street

Artist Name	Danceability	Track
Bleachers	0.824	Hate That You Know Me - MTV Unplugged
Bleachers	0.818	Wake Me
Bleachers	0.814	Shadow
Taylor Swift	0.811	Paper Rings
The Chicks	0.805	Julianna Calm Down
Lorde	0.796	Sober
The Chicks	0.784	Texas Man
Bleachers	0.776	Hate That You Know Me
St. Vincent	0.774	Daddy's Home
Bleachers	0.769	All My Heroes
Taylor Swift	0.766	Look What You Made Me Do
Taylor Swift	0.751	Question...?
Taylor Swift	0.743	Lavender Haze
Taylor Swift	0.739	False God
Bleachers	0.732	Goodbye

Table 3 shows us the 20 songs in Jack Antonoff's discography with the highest `danceability` scores. Bleachers takes up 35% of this list, and Taylor Swift takes up 40% of this list. St Vincent, The Chicks, Lorde, and Florence and the Machine also haves songs up there.

5 Random Forest

Random forests (RF) build an ensemble of decision trees during the training phase, and then use majority voting for classification tasks. Each tree in the ensemble is trained on a random sample of data taken with replacement, and only a random subset of features are considered when splitting at each node. This method ensures that each tree in the ensemble is unique, avoiding overfitting and capturing non-linear relationships and more complex patterns in the data.

5.1 Data Preparation

The categorical feature `producer`, which contains the names of the producer for each track, was one-hot encoded to create binary columns for each producer (i.e. `is_antonoff`, `is_elworth`, etc.). The producer feature was then converted into a factor variable, and the dataset was split into training and tests sets.
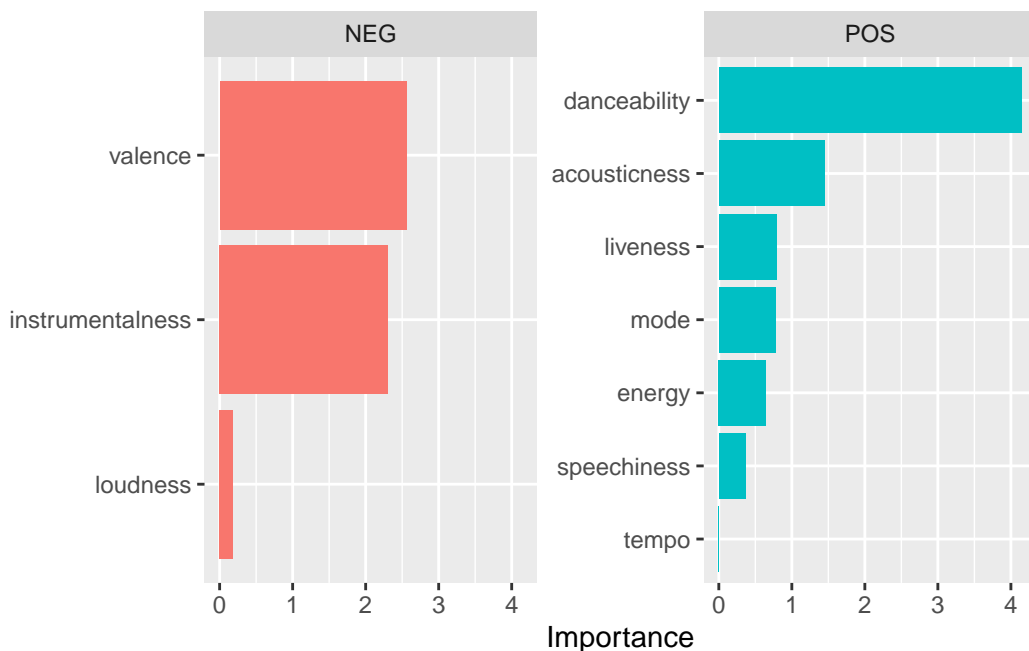
Figure 2: Importance Scores of Each Variable in Determining Antonoff's tracks

5.2 Model Training

The RF was trained using the (**ranger?**) package. The 'importance' parameter was set to 'impurity', which sets the model to interpret variable importance using Gini impurity.

Gini impurity is a measure of misclassification. At every split in a decision tree, the algorithm tests how well each feature splits the data. The features that split the data the most accurately, and then with the least impurity, would have the lowest impurity score. In Random Forests, a variable's importance can be calculated by looking at how much the tree nodes that use that variable reduce impurity on average.

Gini impurity was chosen on the grounds that this study is interested in not only classifying producers, but, more importantly, in understanding how the underlying data informs these decisions. By calculating the Gini impurity not only for the entire model, but specifically for each producer, we can better understand how the variables in our dataset describe the differences between producers. Without song structure, instruments, sound, or other more ephemeral components of style, these differences can be taken as a proxy for style.

5.3 Results

5.3.1 Model Performance

Table 4: Overall Performance Statistics

Metric	Value
Accuracy	0.6111
95% Confidence Interval	(0.579, 0.6425)
No Information Rate	0.4274
P-Value (Acc > NIR)	<2.2e-16
Kappa	0.4223

The accuracy of the model (0.6111) being significantly higher than the No Information Rate (0.4274) means that it is doing more than just blind guessing based on class prevalence. With a p-value below 0.05, we can confirm the model's statistical significance over a naive approach.

The model's accuracy and Kappa statistics suggest a moderate level of performance. However, there is enough demonstrated statistical significance to suggest that there are underlying patterns in the data that differentiate producers.

5.3.2 Stylistic Difference

Correlation heatmaps can help us better understand the differences between each producer's "style". The variables with the most variance were danceability, loudness, and energy.

```
# A tibble: 9 x 4
  variable         min_value max_value range
  <chr>                <dbl>     <dbl> <dbl>
1 energy              -0.779     0.821  1.60
2 loudness            -0.728     0.562  1.29
3 danceability        -0.484     0.553  1.04
4 speechiness         -0.445     0.300  0.745
5 acousticness        -0.449     0.271  0.720
6 mode                -0.295     0.289  0.584
7 instrumentalness    -0.235     0.330  0.565
8 valence             -0.212     0.202  0.414
9 liveness            -0.245     0.0853 0.330
```

Similar to the logistic regression, danceability was also, across all producers, the most important predictor in the classification process.

Variable Importance by Producer
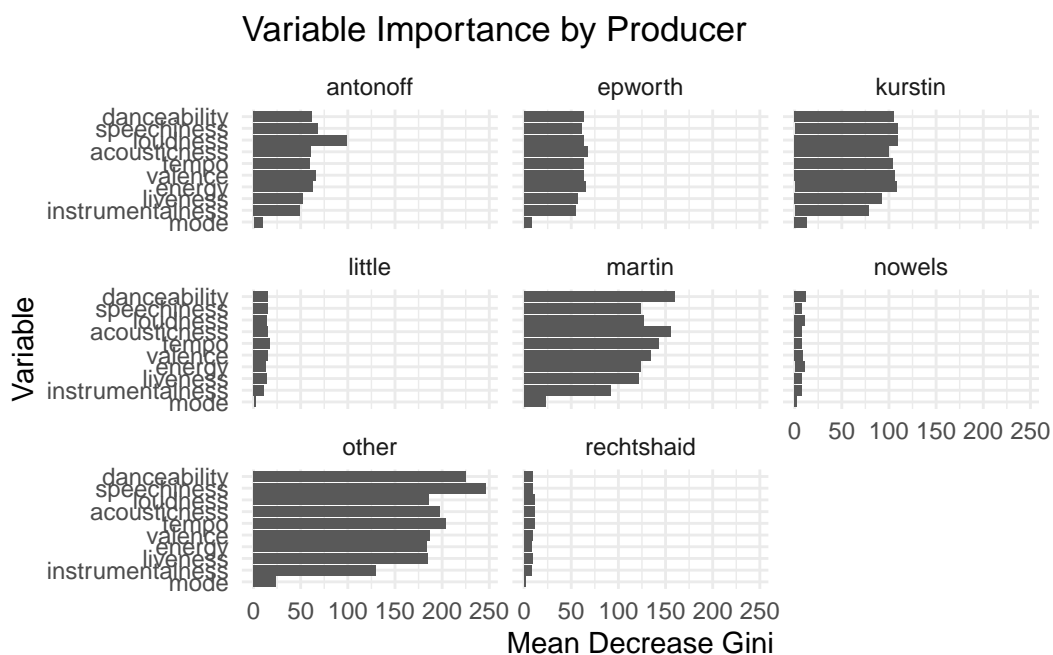
In the visualization of danceability correlations across producers, Antonoff's danceability has the least positive correlation with valence (SCORE), the most negative correlation with acousticness (SCORE), and significant positive correlations with speechiness (SCORE), loudness (SCORE), and energy (SCORE).
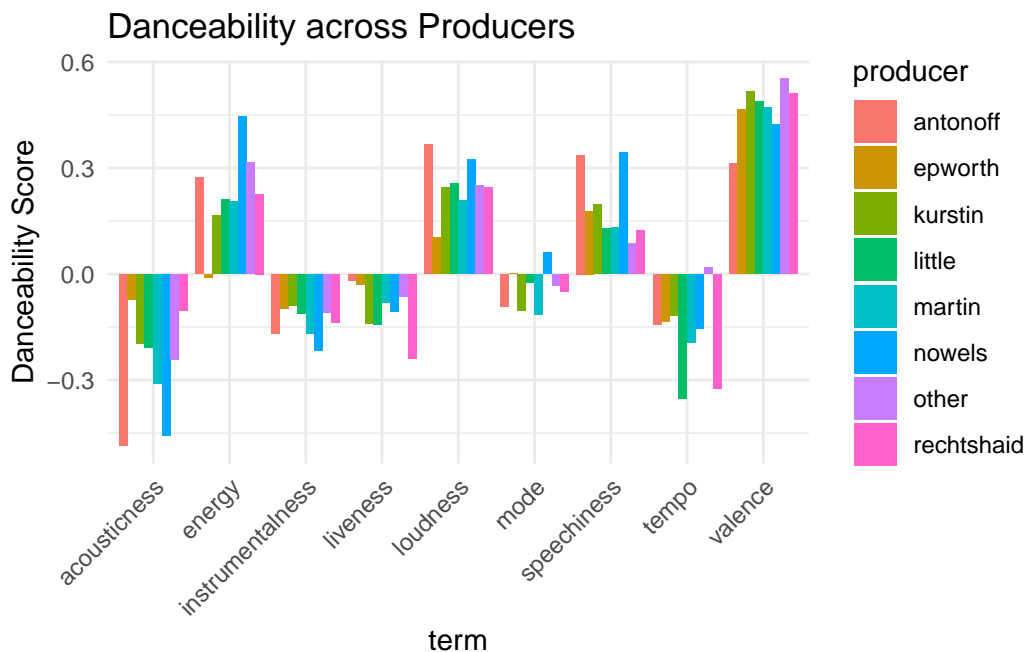
6 Discussion

According to our results, Antonoff's `danceability` score is, compared to other producers, significantly positively correlated with `speechiness`, `loudness` and `energy`, negatively correlated with `acousticness`, and less positively correlated than average with `valence`. Speechy, loud, energetic, synthy, positive-yet-nostalgic. Our model's understanding of Antonoff's style, if we can take what `danceability` 's correlates capture as a proxy for style, tracks with both a close-listen and critics' descriptions of Antonoff's sound.

His tracks have been described as anthemic (cite), with a distinct vocal treatment that makes it sound, as Gamman put it, "the way it would sound to the person who's singing them".

> In the treble mix of the song, like in the very upper end, sort of above where the vocals are – he often really crushes that down. It's lower in volume than anything else, which is sort of a weird effect. When you hear someone speaking, you hear a lot of that, like, noise in their voice. When you yourself are speaking, you hear less of it, right? You hear more of your own voice bouncing around in your head. And so it sort of creates this effect of – **like he mixes his vocals the way it would sound to the person who's singing , them. Which is sort of strange.** He often has no vocals on that upper end, which is very unusual, and then often there are random little bits of noise happening up there which is the sort of thing her likes to do.

Perhaps, without further documentation on the `speechiness` feature's construction, clearing the noise and isolating the middle range of the vocals may make for a stronger `speechiness` feature. Or, more realistically, artists who collaborated with Antonoff happen to be on the more lyrics-oriented end of the spectrum, rather than dance hit pop, similar to Antonoff's own Bleachers and fun..

Turning to the music itself, even Antonoff-produced songs with the highest `danceability` scores aren't exactly dance hits.

While the most danceable song*, I Think He Knows* by Taylor Swift, seems danceable to me, I personally don't hear Taylor Swift's *Vigilante Shit* or Florence and the Machine's *Heaven is Here* as danceable. They aren't understated, ambient tracks either – there's something coercing your body to do something. Like Gamman's description of Antonoff's vocal design sounding like how they would sound to the person who's singing them, these tracks seem like they're trying to get inside of your skin. But they don't have the rhythm of a dance track.

I Wanna Get Better by Bleachers is very similar to the typical Antonoff track that the outputs of our RF describe. This song is more anthemic, closer to the driving-over-the-Brooklyn-bridge, sort of neo-Springsteen emotions that Antonoff is often characterized (or admitting) to striving for, than danceable. It also does seem like it's trying to coerce your attention, if not your body, and the speechiness, loudness, and energy which are distinctly correlated with

danceability in Antonoff's RF results are all apparent. The redemption, broken-but-getting-better-quasi-valence can also be heard in the lyrics, but it still doesn't have the kind of rhythm for dancing.

Cornelia Street by Taylor Swift is another sort of almost-danceable track – while the melody is catchy, the lyrics themselves are wistful. The vocal treatment, again, is there – Swift's vocals feel magnified, carrying the bridge of the song.

What, then, is the `danceability` variable telling us? Antonoff's most "danceable" songs are better understood through their strongest correlates with `danceability`, although `danceability` is the strongest predictor overall.

I am thinking about what Gamman said in his video.

> I did use the phrase "visceral hatred". But i did so only to describe my visceral hatred. I can't stand the sound of it. I hate it. And the way in which I hate it is viscerally. But I didn't go beyond that because it doesn't go beyond that. Nothing against the guy, I like Taylor swift's stuff. It's just subjective taste.

It's just his subjective taste, but Gamman is right. The `danceability` variable, which is the strongest predictor of Jack-ness, seems to actually be more about some sort of "visceral"-ness when you do a close-listen of the songs with the highest danceability scores. Not all of them are danceable, but the one's that aren't still feel like they're trying to manipulate the way your body is holding itself in some way. The reason that Antonoff's sound dominating chart-topping pop production gets under people's skin could just be because his music gets underneath your skin.

References

@calebgamman. 2022. "Im Able to Instantly Detect If Jack Antonoff Worked on a Song Due to a Visceral Hatred of His Production Style." https://twitter.com/calebgamman/status/1583474726797979649.

"Get Track's Audio Features." n.d. https://developer.spotify.com/documentation/web-api/reference/get-audio-features.

Kuhn, Max, and Hadley Wickham. 2020. *Tidymodels: A Collection of Packages for Modeling and Machine Learning Using Tidyverse Principles.* https://www.tidymodels.org.

Marantz, Andrew. 2022. "Jack Antonoff's Pop-Music Collaboration Machine." *The New Yorker*, May. https://www.newyorker.com/magazine/2022/05/23/jack-antonoff-pop-music-collaboration-lorde-taylor-swift.

Silge, Julia. 2023. "To Downsample Imbalanced Data or Not, with #TidyTuesday Bird Feeders." Julia Silge. 2023. https://juliasilge.com/blog/project-feederwatch/.

Thompson, Charlie, Daniel Antal, Josiah Parry, Donal Phipps, and Tom Wolff. 2022. *Spotifyr: R Wrapper for the 'Spotify' Web API.* https://CRAN.R-project.org/package=spotifyr.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wikipedia. 2023. "Jack Antonoff Production Discography — Wikipedia, the Free Encyclopedia." https://en.wikipedia.org/wiki/Jack_Antonoff_production_discography.