Shapes as Product Differentiation:

Neural Network Embedding in the Analysis of Font Markets

Sukjin (Vincent) Han, Eric Schulman,
Kristen Grauman & Santhosh Ramakrishnan

Dept. of Economics & Dept. of Computer Science
U. of Bristol & U. of Texas at Austin

January 3, 2021

Markets for Products with High-Dimensional Attributes

The key attributes in many products considered in economic analyses are unstructured i.e. text/images

Examples:

- design: automobiles, houses, furniture, clothing
- creative works: books, musics and movies

Markets for Products with High-Dimensional Attributes

The key attributes in many products considered in economic analyses are unstructured i.e. text/images

Examples:

- design: automobiles, houses, furniture, clothing
- creative works: books, musics and movies

More generally, products are often presented to consumers in visual/textual forms (along with structured attributes)

- packages in supermarket
- catalogs in e-commerce (e.g., Amazon, Airbnb, Yelp)

Markets for Products with High-Dimensional Attributes

The key attributes in many products considered in economic analyses are unstructured i.e. text/images

Examples:

- design: automobiles, houses, furniture, clothing
- creative works: books, musics and movies

More generally, products are often presented to consumers in visual/textual forms (along with structured attributes)

- packages in supermarket
- catalogs in e-commerce (e.g., Amazon, Airbnb, Yelp)

Important decision factors for consumers and decision variables for producers

Traditional Economic Models

Economic models with product attributes

- Discrete-choice models (McFadden (1973), Berry, Levinsohn & Pakes (1995))
- Hedonic models (Rosen (1974), Bajari & Benkard (2005))

These models include

- low-dim observed attributes

  and

- scalar unobserved attributes

Traditional Economic Models

Economic models with product attributes

- Discrete-choice models (McFadden (1973), Berry, Levinsohn & Pakes (1995))
- Hedonic models (Rosen (1974), Bajari & Benkard (2005))

These models include

- low-dim observed attributes

  and

- scalar unobserved attributes

Q: text/images as scalar unobservable vs. high-dim observables?

- may depend on policy questions

Possible Policy Questions

...that can be answered by considering high-dim, unstructured attributes:

- mergers and product variety
- policies that protects the originality of artistic features
- evolution of style (e.g., fashion trend)

Possible Policy Questions

...that can be answered by considering high-dim, unstructured attributes:

- mergers and product variety
- policies that protects the originality of artistic features
- evolution of style (e.g., fashion trend)

This Project: Markets for Fonts

Considers unstructured attributes in particular design product: fonts

This Project: Markets for Fonts

Considers unstructured attributes in particular design product: fonts

Why font market?

1. unstructed visual info on shape helps predict functionality and value of the product (unlike fine art products)

This Project: Markets for Fonts

Considers unstructured attributes in particular design product: fonts

Why font market?

1. unstructed visual info on shape helps predict functionality and value of the product (unlike fine art products)
2. font shapes are relatively simple among visually-differentiated products
   - typefaces are 2-dim, monochrome, a fixed number of characters and mostly describe the product

This Project: Markets for Fonts

Considers unstructured attributes in particular design product: fonts

Why font market?

1. unstructed visual info on shape helps predict functionality and value of the product (unlike fine art products)
2. font shapes are relatively simple among visually-differentiated products
   - typefaces are 2-dim, monochrome, a fixed number of characters and mostly describe the product
3. fonts market is typically large (unlike fine art markets)
   - frequent production and transactions
   - e.g., the online marketplace we consider has over 28,000 fonts and 2,400,000 transactions (past six years)

This Project: Markets for Fonts

Considers unstructured attributes in particular design product: fonts

Why font market?

1. unstructed visual info on shape helps predict functionality and value of the product (unlike fine art products)
2. font shapes are relatively simple among visually-differentiated products
   - typefaces are 2-dim, monochrome, a fixed number of characters and mostly describe the product
3. fonts market is typically large (unlike fine art markets)
   - frequent production and transactions
   - e.g., the online marketplace we consider has over 28,000 fonts and 2,400,000 transactions (past six years)
4. fonts are an illustrative example of the market for other design attributes

Embedding and Low-Dimensional Product Space

We represent font shapes as low-dim number of attributes (like a product space) by neural network embedding

Embedding and Low-Dimensional Product Space

We represent font shapes as low-dim number of attributes (like a product space) by neural network embedding

Specifically, we adapt a state-of-the-art method in deep convolutional neural network

- the network directly learns the embedding,
- i.e., a mapping from font images to a compact Euclidean space
- The embeddings can be thought of as a product space

Embedding and Low-Dimensional Product Space

We represent font shapes as low-dim number of attributes (like a product space) by neural network embedding

Specifically, we adapt a state-of-the-art method in deep convolutional neural network

- the network directly learns the embedding,
- i.e., a mapping from font images to a compact Euclidean space
- The embeddings can be thought of as a product space

Post-dim-reduction verification

- Using tags (i.e., descriptive phrases) assigned to each font by font designers and consumers
- High degree of mutual information between tags and image embeddings

Embedding and Low-Dimensional Product Space

Why convolutional neural network (CNN)?

- for visual/textual data considered in this project, NN outperforms (by large margin) other machine learning methods (e.g., LASSO, random forest)
- esp. CNN is known to be appropriate to capture the "spatial" dependence, e.g., of pixels or musical notes

Embedding and Low-Dimensional Product Space

Why convolutional neural network (CNN)?

- for visual/textual data considered in this project, NN outperforms (by large margin) other machine learning methods (e.g., LASSO, random forest)
- esp. CNN is known to be appropriate to capture the "spatial" dependence, e.g., of pixels or musical notes

Drawback of NN: interpretability

Still, distance between the embeddings has a clear interpretation of visual similarity and can help answer policy questions

Economic Analyses Using Embedding

We conduct two analyses:

1. a simple trend analysis of font style
2. a causal analysis of merger effects on product differentiation using synthetic control method

Economic Analyses Using Embedding

We conduct two analyses:

1. a simple trend analysis of font style
2. a causal analysis of merger effects on product differentiation using synthetic control method

Can think about product differentiation as the outcome of Hotelling-type of spatial competition

Contributions of This Paper

1. Use of embeddings to analyze markets with text/visual data
    - Hoberg & Phillips (2016), Gross (2016), Glaeser et al. (2018), Kozlowski et al. (2019), Gentzkow et al. (2019)

Contributions of This Paper

1. Use of embeddings to analyze markets with text/visual data
   - Hoberg & Phillips (2016), Gross (2016), Glaeser et al. (2018), Kozlowski et al. (2019), Gentzkow et al. (2019)
2. Mergers and product differentiation
   - Berry & Waldfogel (2001), Sweeting (2013)
   - unlike them, we focus on unstructured attributes and follow program evaluation approach

Contributions of This Paper

1. Use of embeddings to analyze markets with text/visual data
   - Hoberg & Phillips (2016), Gross (2016), Glaeser et al. (2018), Kozlowski et al. (2019), Gentzkow et al. (2019)
2. Mergers and product differentiation
   - Berry & Waldfogel (2001), Sweeting (2013)
   - unlike them, we focus on unstructured attributes and follow program evaluation approach
3. Visual attributes as endogenous decision
   - Gillen et al. (2015), Chernozhukov et al. (2017)

Contributions of This Paper

1. Use of embeddings to analyze markets with text/visual data
   - Hoberg & Phillips (2016), Gross (2016), Glaeser et al. (2018), Kozlowski et al. (2019), Gentzkow et al. (2019)
2. Mergers and product differentiation
   - Berry & Waldfogel (2001), Sweeting (2013)
   - unlike them, we focus on unstructured attributes and follow program evaluation approach
3. Visual attributes as endogenous decision
   - Gillen et al. (2015), Chernozhukov et al. (2017)
4. Predictive/trend analysis of design product and visual attributes
   - Font: O'Donovan et al. (2014), Campbell & Kautz (2014)
   - Cloths, furniture and cars: Al-Halah et al. (2017), Mall et al. (2019), Yu & Grauman (2019), Burnap et al. (2016) and Dosovitskiy et al. (2016)

Contributions of This Paper

1. Use of embeddings to analyze markets with text/visual data
   - Hoberg & Phillips (2016), Gross (2016), Glaeser et al. (2018), Kozlowski et al. (2019), Gentzkow et al. (2019)
2. Mergers and product differentiation
   - Berry & Waldfogel (2001), Sweeting (2013)
   - unlike them, we focus on unstructured attributes and follow program evaluation approach
3. Visual attributes as endogenous decision
   - Gillen et al. (2015), Chernozhukov et al. (2017)
4. Predictive/trend analysis of design product and visual attributes
   - Font: O'Donovan et al. (2014), Campbell & Kautz (2014)
   - Cloths, furniture and cars: Al-Halah et al. (2017), Mall et al. (2019), Yu & Grauman (2019), Burnap et al. (2016) and Dosovitskiy et al. (2016)

Roadmap

I. Online marketplace for fonts

II. Construction of embedding and product space

III. Economic analyses

IV. Conclusions

I. Online Marketplace for Fonts

Background: Online Marketplace for Fonts

We consider the world's largest online market place "MyFonts.com" that sell around 28,000 different fonts

- owned by Monotype, sells fonts designed by Monotype
- fonts from third parties foundries

Font are sold as software

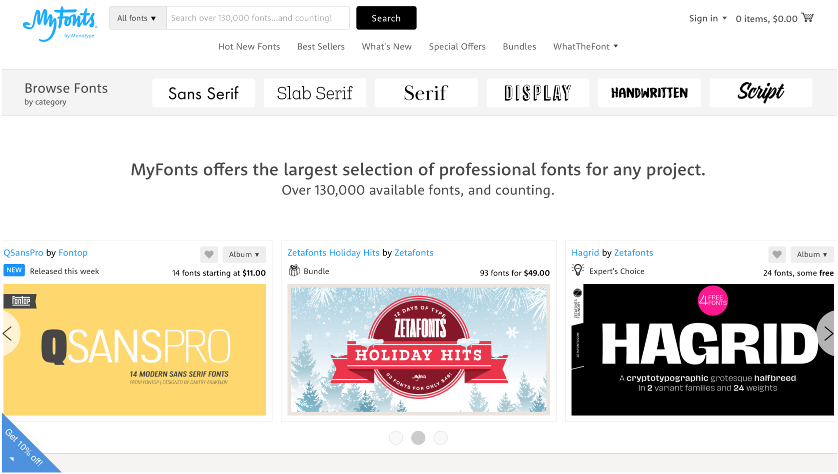
- fonts deliver typefaces
- sold as licenses are protected by the End User License Agreement (EULA)

In this market, consumers are typically other designers who use fonts as intermediate goods to produce...

- prints (posters, pamphlets, cards)—desktop license
- webpages—web license

Background: Online Marketplace for Fonts

Main page in MyFonts.com



Background: Online Marketplace for Fonts

Individual fonts are called styles, groups of styles are called families;
Example of a font family page in MyFonts.com

Gilroy Light Italic *Gilroy Light Italic*	from **$25.00**	Buying Choices
Gilroy Regular Gilroy Regular	from **$25.00**	Buying Choices
Gilroy Regular Italic *Gilroy Regular Italic*	from **$25.00**	Buying Choices
Gilroy Medium Gilroy Medium	from **$25.00**	Buying Choices
Gilroy Medium Italic *Gilroy Medium Italic*	from **$25.00**	Buying Choices
Gilroy Semi Bold **Gilroy Semi Bold**	from **$25.00**	Buying Choices

Data from the Marketplace

Sample between 2000-2017

- number of fonts: 28,659
- number of orders (every minute): 2,446,604

Product attributes

- Unstructured
  - images of typefaces
  - tags (descriptive words assigned by producers or consumers)
- Structured
  - price, foundry/designer info, date introduced
  - license type (desktop, web, apps, ePub, digital ads)
  - number of languages/glyphs supported

Visual Attributes

Fonts are displayed on the webpage using pangrams

- effectively capture important design elements (spacing, deep-height, up-height, ligature)

Format of pangram images: bitmap (200×1000 pixels)

We use (crops of) pangrams as direct inputs in CNN

- The neural network is designed to classify font styles
- Crops create multiple instances of the same style

Examples of Pangram Images

Quick zephyrs blow, vexing daft Jim.

Quick zephyrs blow, vexing daft Jim.

Quick zephyrs blow, vexing daft Jim.

Quick zephyrs blow, vexing daft Jim.

Quick zephyrs blow, vexing daft Jim.

Quick zephyrs blow, vexing daft Jim.

II. Construction of Embedding and Product Space

Construction of Embedding

We employ a method where the network directly learns a mapping from pangram images to a compact Euclidean space

- this mapping is called embedding
- we map each pangram to 128-dim embedding
- L^2 distance corresponds to measure of similarity of font shape

Developed by Schroff et al. (2015) for face recognition

- classification by thresholding distance between embeddings
- perform substantially better than earlier classification network

Construction of Embedding

We employ a method where the network directly learns a mapping from pangram images to a compact Euclidean space

- this mapping is called embedding
- we map each pangram to 128-dim embedding
- L^2 distance corresponds to measure of similarity of font shape

Developed by Schroff et al. (2015) for face recognition

- classification by thresholding distance between embeddings
- perform substantially better than earlier classification network

We adapt their approach for our purpose

- not interested in classification of font identity
- but embedding and resulting product space is our interest

Triplet Loss

Triplet (Weinberger et al. (2006), Wang et al. (2014))

- triplet i: anchor x_i^a, positive x_i^p, negative x_i^n

$$||f(x_i^a) - f(x_i^p)||_2^2 + \alpha \leq ||f(x_i^a) - f(x_i^n)||_2^2 \qquad (1)$$

$\forall (x_i^a, x_i^p, x_i^n) \in T$, where α is an enforced margin

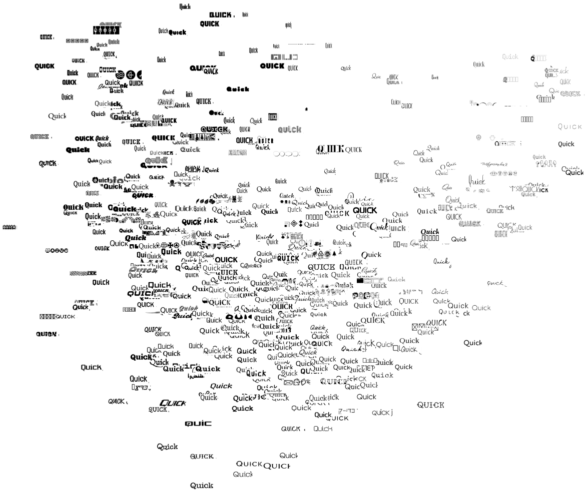
Triplet-based loss function that is minimized is

$$L = \sum_i^N [||f(x_i^a) - f(x_i^p)||_2^2 - ||f(x_i^a) - f(x_i^n)||_2^2 + \alpha]_+$$

Constructed Product Space

128-dim product space, projected in 2-dim for visualization

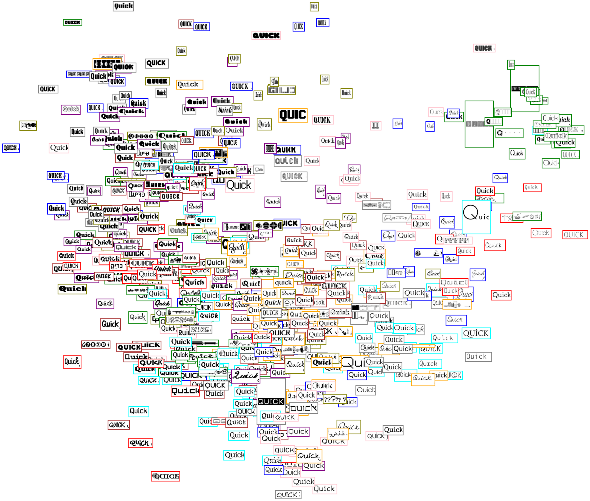
- each point corresponds to embedding of each font family



Constructed Product Space

128-dim product space, projected in 2-dim for visualization

- each point corresponds to embedding of each font family



Internal Evaluation

Test (validation) set: positive (same) and negative (different) pairs

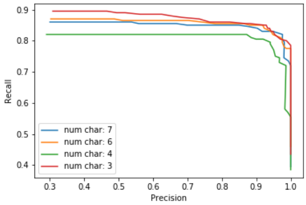
- from the embedding, obtain true accepts and false accepts

Statistics:

- Accuracy (TP+TN)/(P+N): 0.8925 ± 0.01083
- Validation rate (true acceptance rate TP/(TP+FN)) and false acceptance rate FP/(TN+FP):

$$VAL = 0.09375 \pm 0.01875 \text{ at } FAR = 0.0$$

- Precision TP/(TP+FP) and recall TP/(TP+FN) curve



External Verification Using Word Embedding

Want to verify that visual attributes captured in the resulting embedding are relevant to economic agents' perception

External Verification Using Word Embedding

Want to verify that visual attributes captured in the resulting embedding are relevant to economic agents' perception

Measure correlation with "perceived" attributes

- tags assigned to each font family by font designers and consumers
- also high-dim: nearly 30,000 different words in the tags
  - e.g., curly, flowing, geometric, organic, decorative, contrast
- apply standard word embedding "Word2vec" (2-layer NN)

Compare k-mean clusterings (with 60 clusters) from shape embeddings and word embeddings

- measure how well they match using mutual information

Mutual Information

How well these two sets of clusters match?

Normalized mutual information (NMI):

$$NMI(F, W) = \frac{I(F, W)}{[H(F) + H(W)]/2}$$

where $H(\cdot)$ is entropy and $I(F, W) = H(F) - H(F|W)$ is mutual info between F and W

- how much informative W is in determining F
- value between 0 and 1 (0 being W contains no info for F)

Mutual Information

How well these two sets of clusters match?

Normalized mutual information (NMI):

$$NMI(F, W) = \frac{I(F, W)}{[H(F) + H(W)]/2}$$

where $H(\cdot)$ is entropy and $I(F, W) = H(F) - H(F|W)$ is mutual info between F and W

- how much informative W is in determining F
- value between 0 and 1 (0 being W contains no info for F)

We obtain $NMI(F, W) = 0.484$, which is quite promising

- cf. NMI between industry-defined product category and W $= 0.261$

III. Economic Analyses

1. Trend Analysis

As a reduced-form analysis, using shape embeddings, we analyze the supply- and demand-side trends in font style

Supply-side

- there are constant entry of new products in the marketplace

Demand-side

- there are on average more than 1,000 fonts (stably) sold per day

Trends in Font Style: Summary of Findings

1. Fonts become more innovative over time

2. Different preference over shapes, depending on license type

   - more conservative shapes for web license (e.g., webpages)
   - more adventurous shapes for desktop license (e.g., printed materials)

2. Effect of Merger on Product Differentiation

How did product differentiation change after the merger between FontFont and Monotype in July 2014?

2. Effect of Merger on Product Differentiation

How did product differentiation change after the merger between FontFont and Monotype in July 2014?

Two measures of visual differentiation using the embeddings:

1. distance from benchmark font Averia: $D_i \equiv \|f(x_i) - f_{averia}\|_2$
2. "gravity" measure:

$$\tilde{D}_i \equiv -\sum_{j \neq i} \frac{1}{\|f(x_i) - f(x_j)\|_2}$$

We take average of D_i or $\tilde{D}_i$ of all new fonts created by a foundry in given period

- considering two measures of differentiation provides a robustness check
- "gravity" measure does not depend on the benchmark

Effect of Merger on Product Differentiation

Want to estimate how FontFont's design decision has changed by the merger

- note: even before the merger, FontFont has sold its fonts in MyFonts.com as a third party

Effect of Merger on Product Differentiation

Want to estimate how FontFont's design decision has changed by the merger

- note: even before the merger, FontFont has sold its fonts in MyFonts.com as a third party

Challenges:

- only single treated unit (FontFont), multiple untreated (control) units
- difficult to find a single control unit that matches treated unit

Synthetic Control Method

Synthetic control method addresses these challenges:

- Abadie & Gardeazabal (2003), Abadie et al. (2010)

Compare treated unit with a "synthetic control unit"...

- = a weighted average of all control units

Synthetic Control Method

Synthetic control method addresses these challenges:

- Abadie & Gardeazabal (2003), Abadie et al. (2010)

Compare treated unit with a "synthetic control unit"...

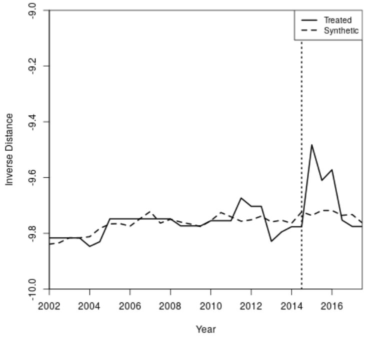
- = a weighted average of all control units

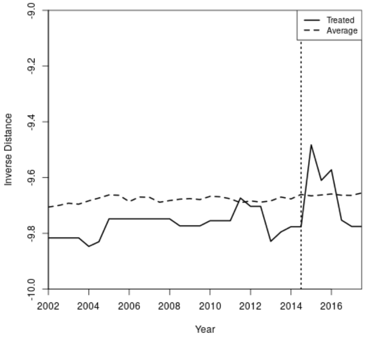
Weights W: estimated by minimizing $\|X_1 - \boldsymbol{X}_0 W\|$

- vector X_1: treated unit's observed characteristics (including pre-trt outcomes)
- matrix $\boldsymbol{X}_0$: control units' characteristics
- we use: embeddings (pre-treatment), glyph count, sales, order count, age

Treated Unit vs. Synthetic Control

Trends of treated unit vs. synthetic control (left), vs. naive control avg. (right)





Estimated Treatment Effects

Years (After Merger)	2015/1–2015/2	2016/1–2016/2	2017/1–2017/2
Treatment Effect	0.1070	0.0576	-0.019
p-value (block)	0.0370	0.0741	1
p-value (all)	0.0022	0.0522	0.9978

- *p*-values are computed based on Chernozhukov et al. (2019)
- also conducted placebo test before the merger

FontFont creates more experimental fonts (i.e., increases product variety), at least temporarily

Estimated Treatment Effects

Years (After Merger)	2015/1–2015/2	2016/1–2016/2	2017/1–2017/2
Treatment Effect	0.1070	0.0576	-0.019
p-value (block)	0.0370	0.0741	1
p-value (all)	0.0022	0.0522	0.9978

- *p*-values are computed based on Chernozhukov et al. (2019)
- also conducted placebo test before the merger

FontFont creates more experimental fonts (i.e., increases product variety), at least temporarily

Other tranditional measures of differentiation (e.g. glyph counts, # of new fonts) do *not* capture this merger effect

Estimated Treatment Effects

Possible reasons for the significant merger effect:

1. increases visual variety to diversify, as merger promotes efficiency
2. avoids cannibalization, i.e., competition of their own

Conclusions

Summary

Consider simplest design product, fonts, and quantify shapes using deep neural network embedding

The resulting low-dim product space can be a basis for various economic analyses

- its distance measures product similarity

Conduct two economic analyses using the embeddings

- trend analysis of font style
- merger analysis with causal interpretation using synthetic control method

Thank You!!