face of
fonts
字之魂

face of

fonts

字之魂

中国梦

If you want to make a superhero movie,

What font will you choose?

THE AVENGERS

THE AVENGERS

THE AVENGERS


THE AVENGERS

字之魂

A font can make or break your **brand**

A typographic designer
must convey **meaning**
through visuals

字之魂

字之魂

image classification to fonts

+

CNNs for NLP

data-driven approach to
augment the process

字之魂

face of fonts | 字之魂

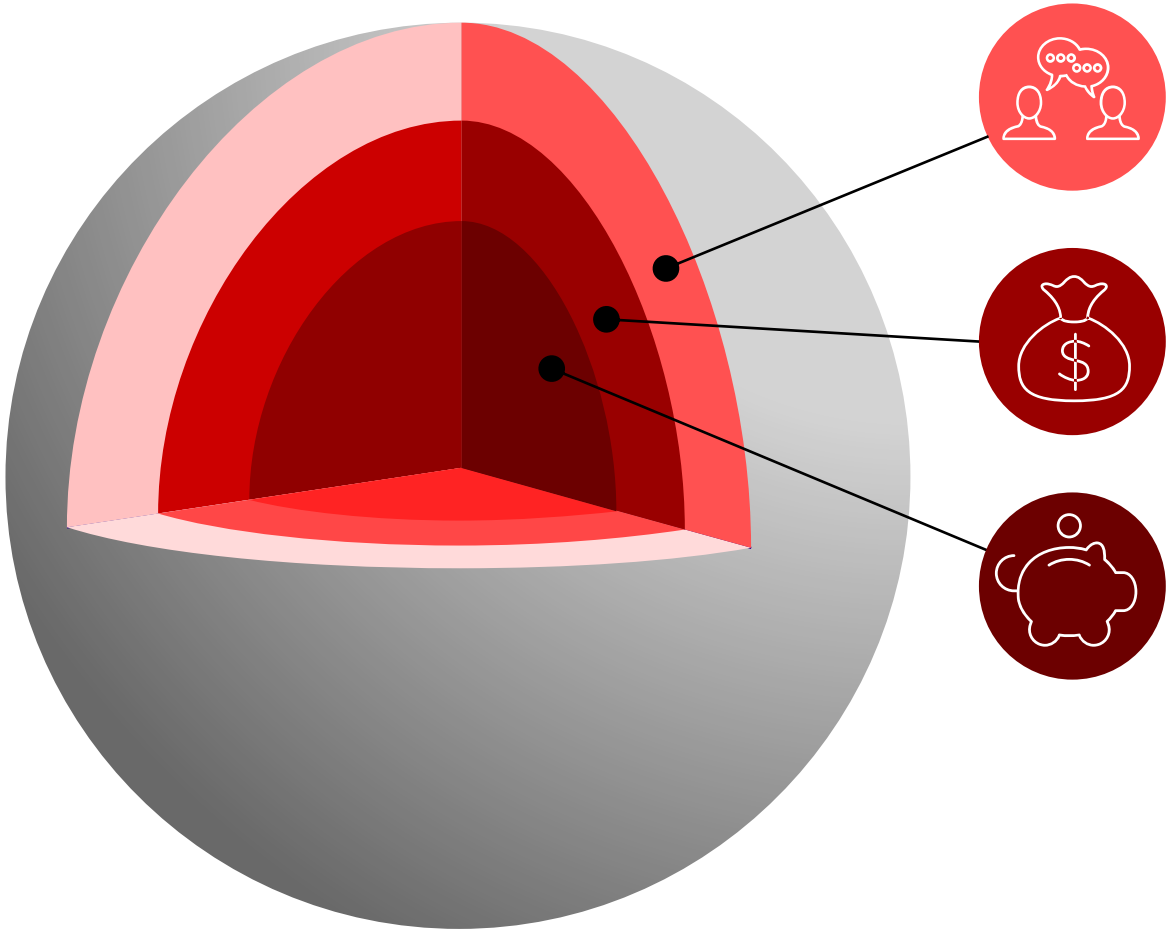
First cut for designers

New modality for NLP

Convergence in available tools

Large free database of fonts (GoogleFonts)

Objective



Word Usage

Natural Language processing of sentences

Word Form

Image classification of font so we can manipulate it

Extract semantic meaning

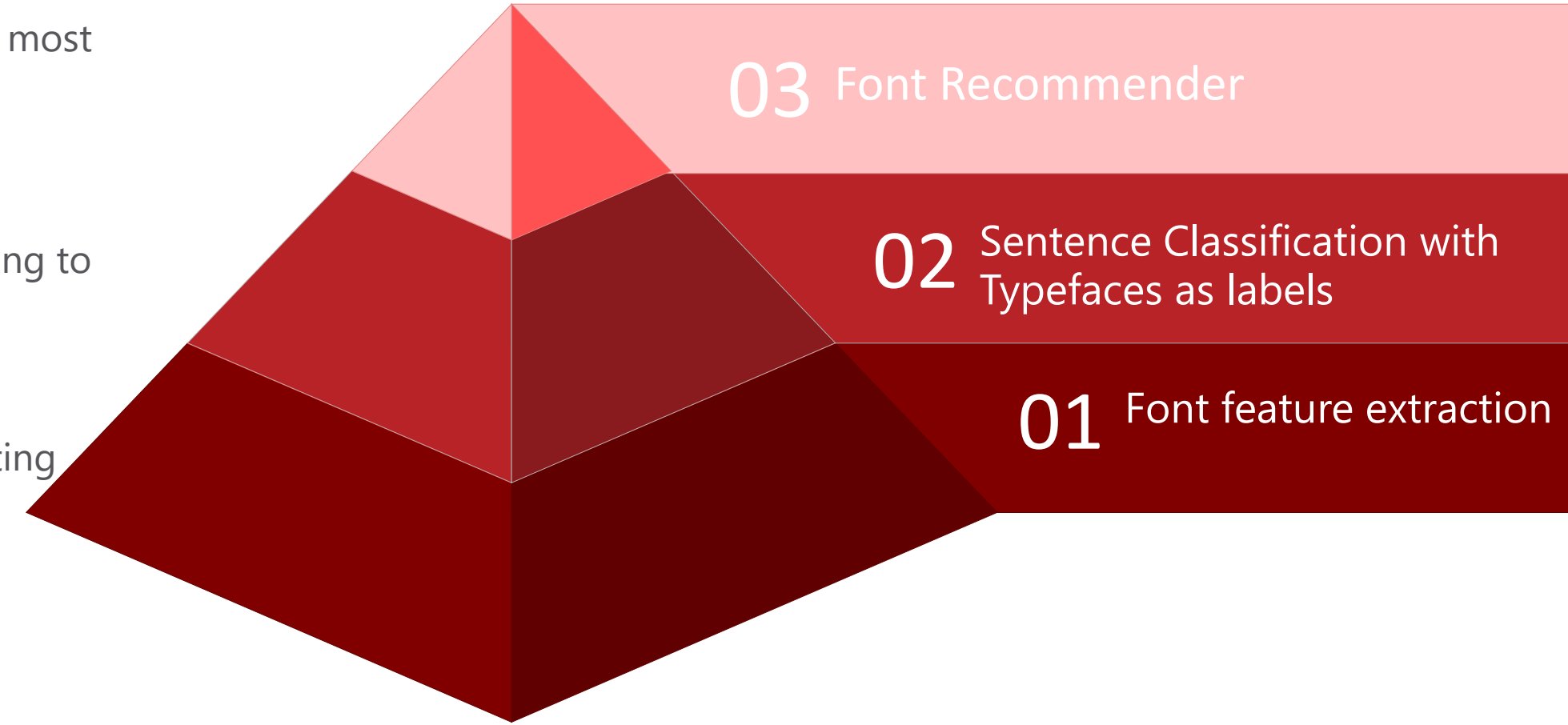
Recommend appropriate font from unseen data

Methodology

3 Model to predict most appropriate font

2 Able to classify typefaces according to usage

1 Vectors representing fonts

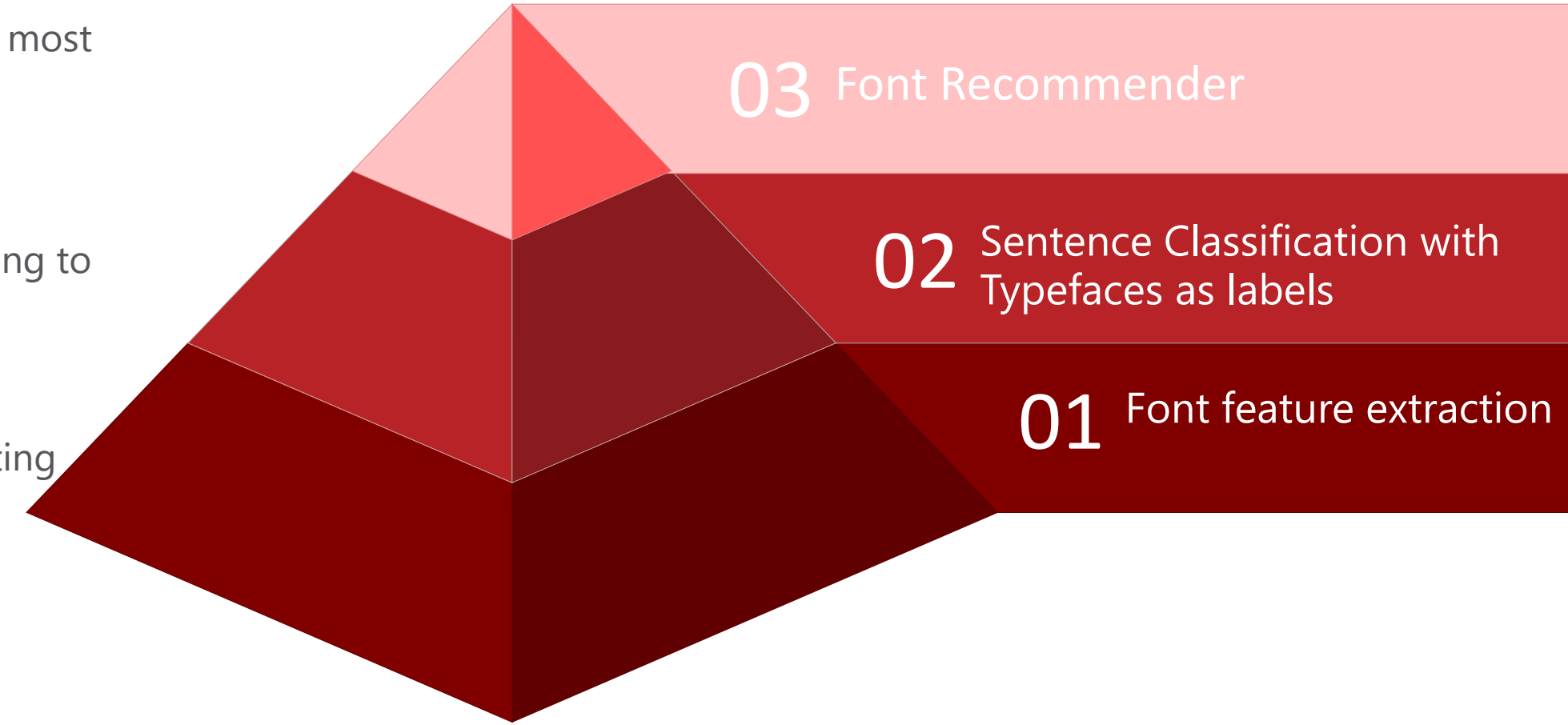


Methodology

3 Model to predict most appropriate font

2 Able to classify typefaces according to usage

1 Vectors representing fonts

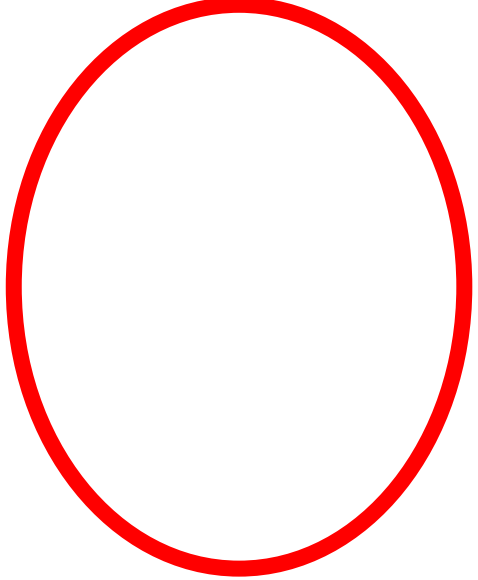


Fontjoy

1883 different fonts from GoogleFonts

Font image to font vector

Allow you to manipulate fonts



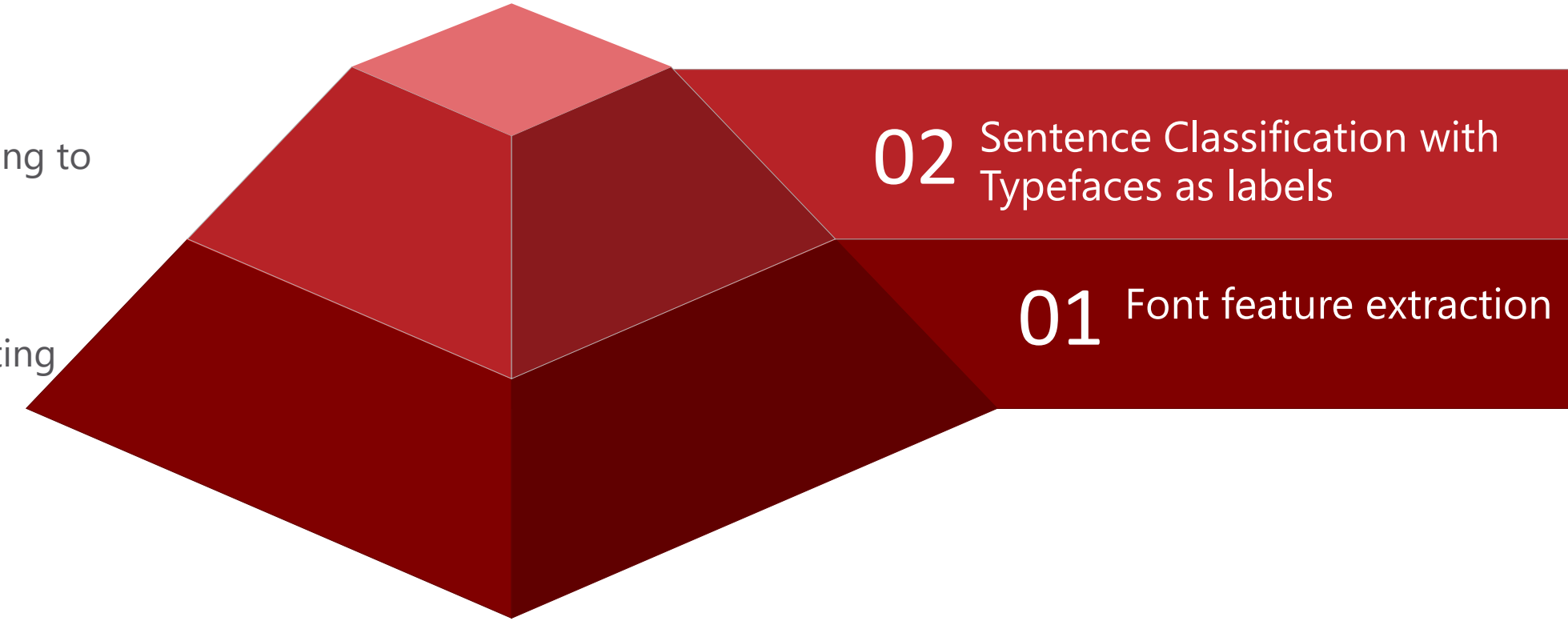
Can identify clusters of similar fonts

Qiao,J (2017)

Methodology

2 Able to classify typefaces according to usage

1 Vectors representing fonts



workflow

Must be from the same source

Example: LinkedIn

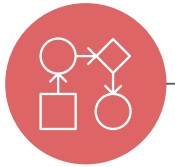
Dataset
Taken from LinkedIn

Pre-processing
Convert dataset into usable format

Natural Language Analysis with Python NLTK

Clean up dataset

Hypothesis: Typography bears semantic meaning



LIBLINEAR
Baseline result to test hypothesis

CNN
To train a usable model

Proven to be effective in sentence classification

workflow

Must be from the same source

Example: LinkedIn

Dataset
Taken from LinkedIn

Pre-processing
Convert dataset into usable format

Natural Language Analysis with Python NLTK

Clean up dataset

Hypothesis: Typography bears semantic meaning

LIBLINEAR
Baseline result to test hypothesis

CNN
To train a usable model

Proven to be effective in sentence classification

workflow

Must be from the same source

Example: LinkedIn

Dataset

Taken from LinkedIn

No available dataset that suited our purpose

workflow

Must be from the same source

Example: LinkedIn

Dataset
Taken from LinkedIn

Criteria:

1. **Font from the same domain is used**
   Meaning is constrained by the different domains that the typeface is used on

2. **Large enough number of instances**
   Sufficient amount of data

3. **Publicly available corporate branding guides**
   Official information provided by companies

workflow

LinkedIn

Must be from the same source

Dataset
Taken from LinkedIn

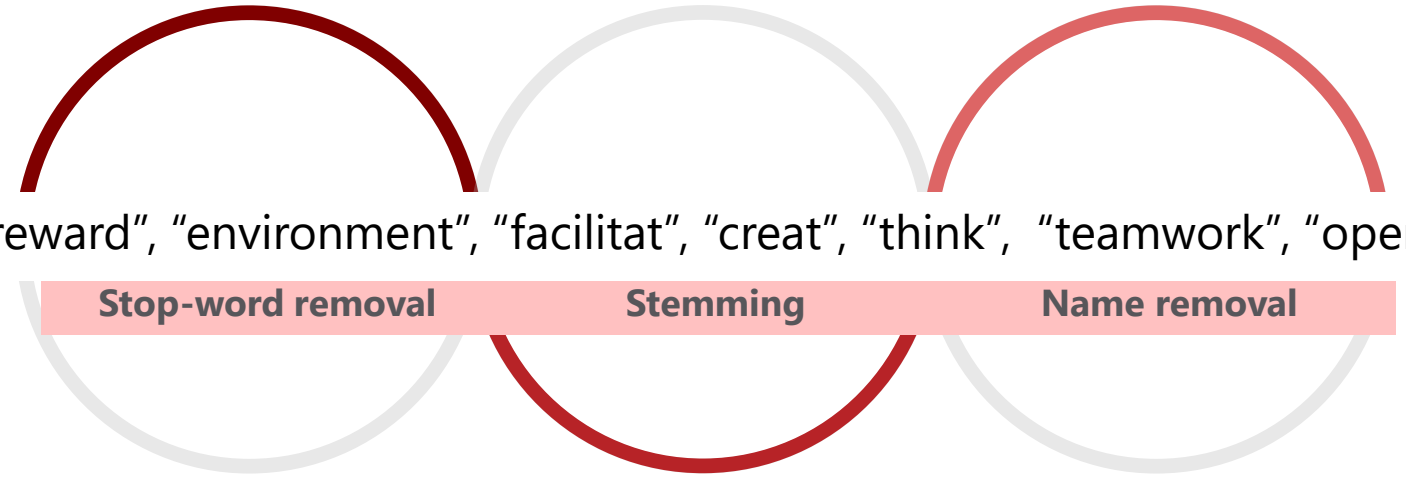
Universities and government agencies

Lifestyle brands

Pre-processing

Convert dataset into usable format

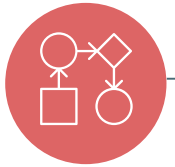
Natural Language Analysis with Python NLTK



P “Provid”, “challeng”, “reward”, “environment”, “facilitat”, “creat”, “think”, “teamwork”, “open”, “communicat” ɔns.

Example Sentence

Hypothesis:
Typography bears semantic meaning



LIBLINEAR

Baseline result to test hypothesis

Chosen due to:

1. **Fast classification time**

Designed to be highly efficient

2. **Able to try multi-class classification**

Easily tune parameters to provide preliminary result

Chang,K , Hsieh, C, Wang X, Fan, R & C, Lin (2017)

CNN

To train a usable model

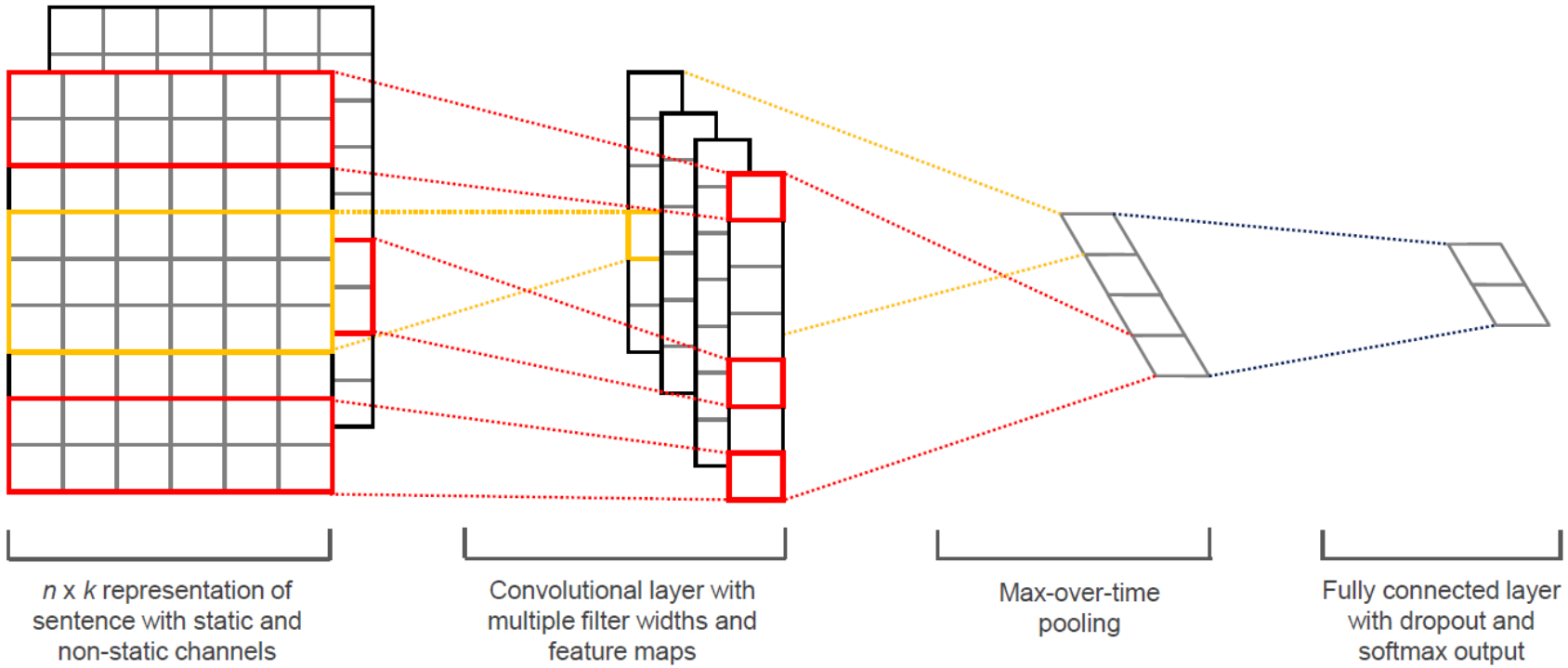
Proven to be effective in sentence classification

Utilize the state-of-the-art in sentence classification using CNN referencing a practitioner's guide

CNN

To train a usable model

Proven to be effective in sentence classification



CNN model inspired by Yoon Kim's paper

Y,Kim (2014)

CNN

To train a usable model

Proven to be effective in sentence classification

Maxlength = 200

nb_filter/ n_gram	1	2	3	4
1000	38.75%	40.53%	39.42%	39.64%
3000	37.86%	39.42 %	37.63%	40.98%
5000	44.58%	45.75%	37.86%	34.96%

The highest accuracy obtained was 45.75%

Therefore we utilized this model to generate font recommendations

workflow

Dataset
Taken from LinkedIn

LinkedIn

Must be from the same source

Pre-processing
Convert dataset into usable format

Natural Language Analysis with Python NLTK

Clean up dataset

LIBLINEAR
Baseline result to test hypothesis

Hypothesis: Typography bears semantic meaning

CNN
To train a usable model

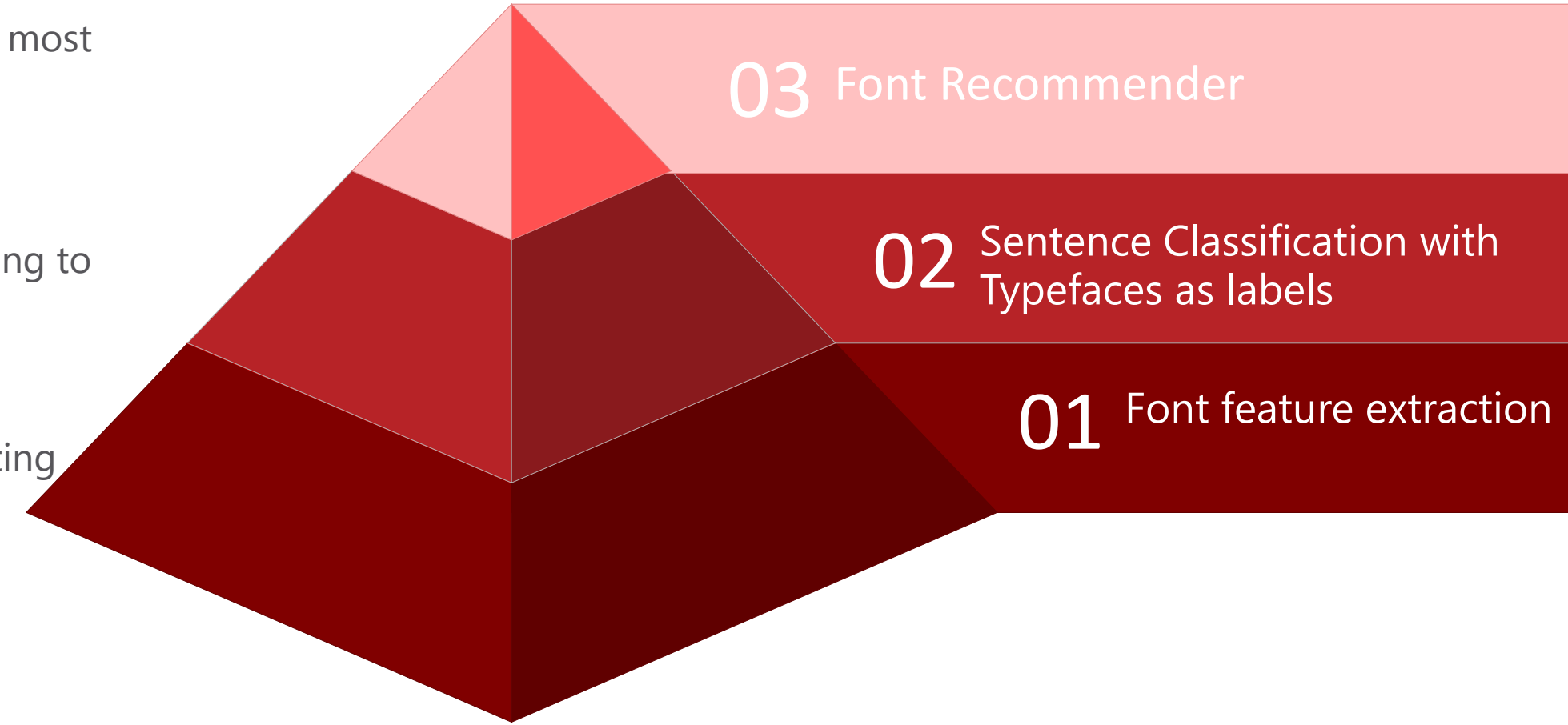
Proven to be effective in sentence classification

Methodology

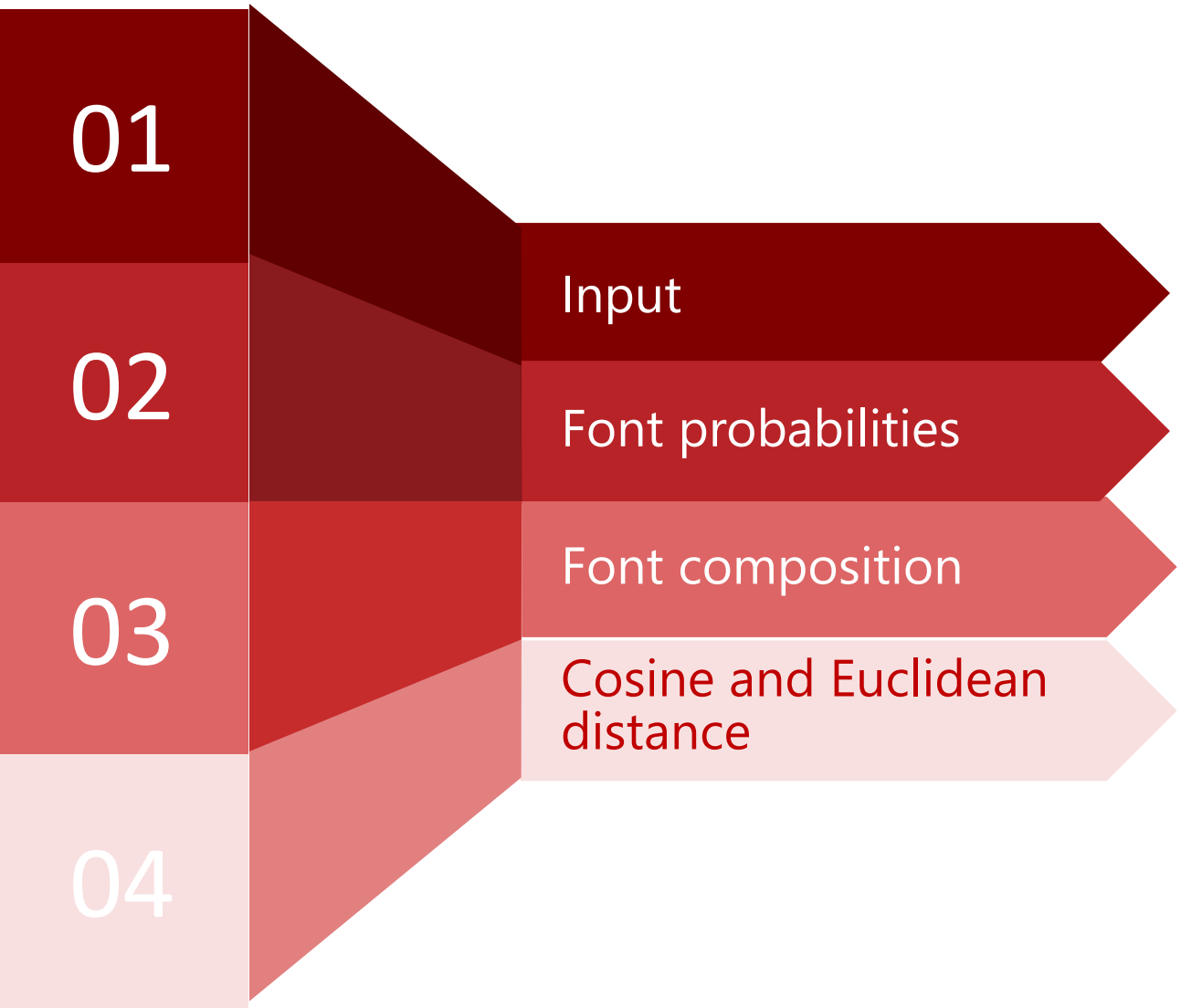
3 Model to predict most appropriate font

2 Able to classify typefaces according to usage

1 Vectors representing fonts



Font Recommender



Font Recommender

01 Input

02 Font probabilities

03 Font composition

04 Cosine and Euclidean distance

Type out a sentence that best represents what you want to convey

"The Palestinian Museum is an independent institution dedicated to supporting an open and dynamic Palestinian culture..."

Font Recommender

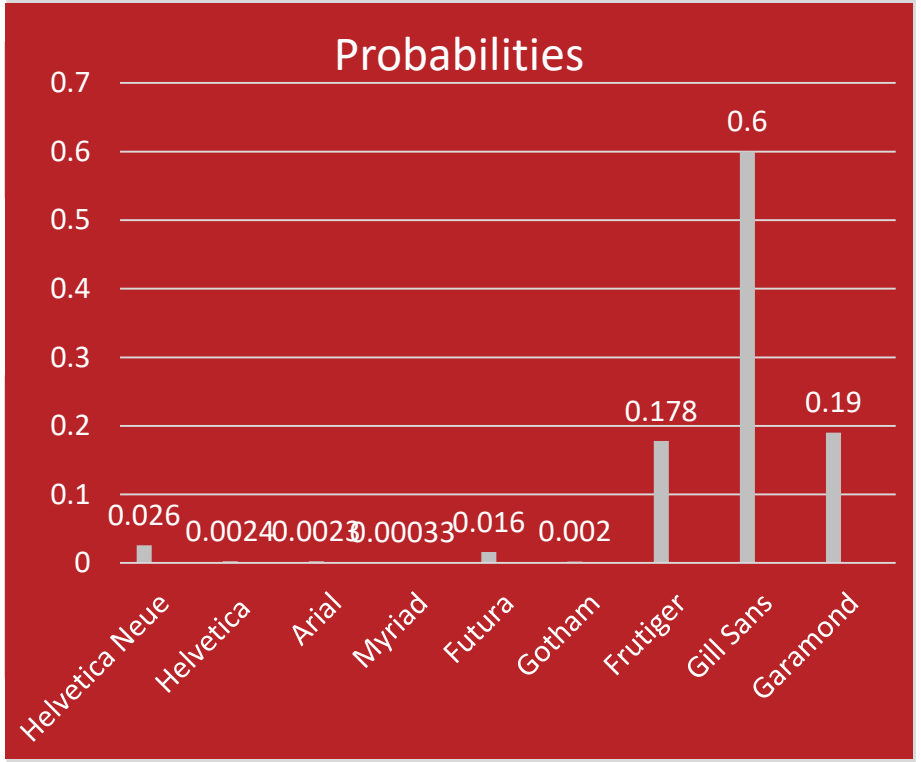
01 Input

02 Font probabilities

03 Font composition

04 Cosine and Euclidean distance

Behind the scenes:



Font Recommender

01 Input

02 Font probabilities

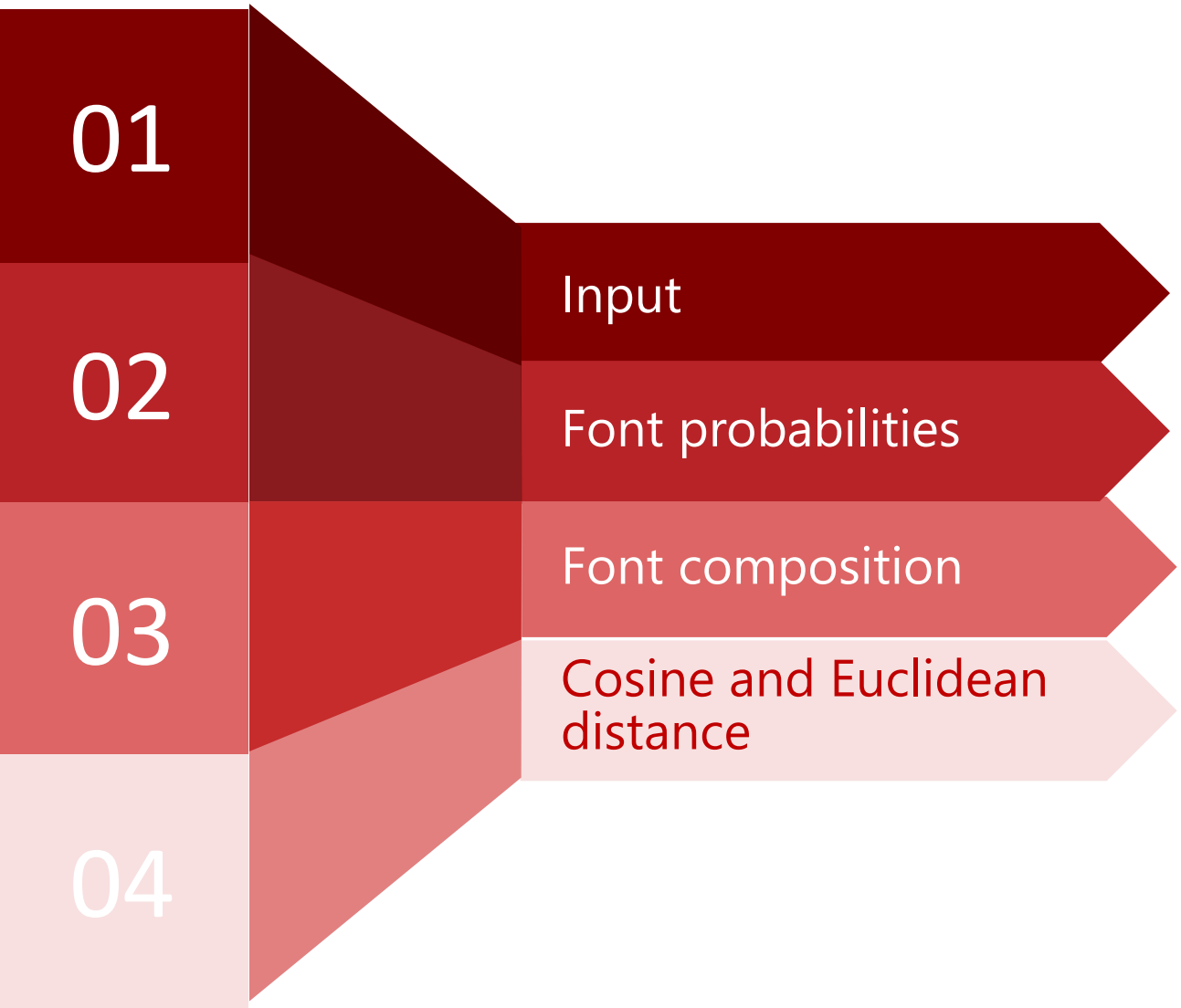
03 Font composition

04 Cosine and Euclidean distance

Behind the scenes:

$$\sum_{i=0}^{n} font_i prob_i$$

Font Recommender



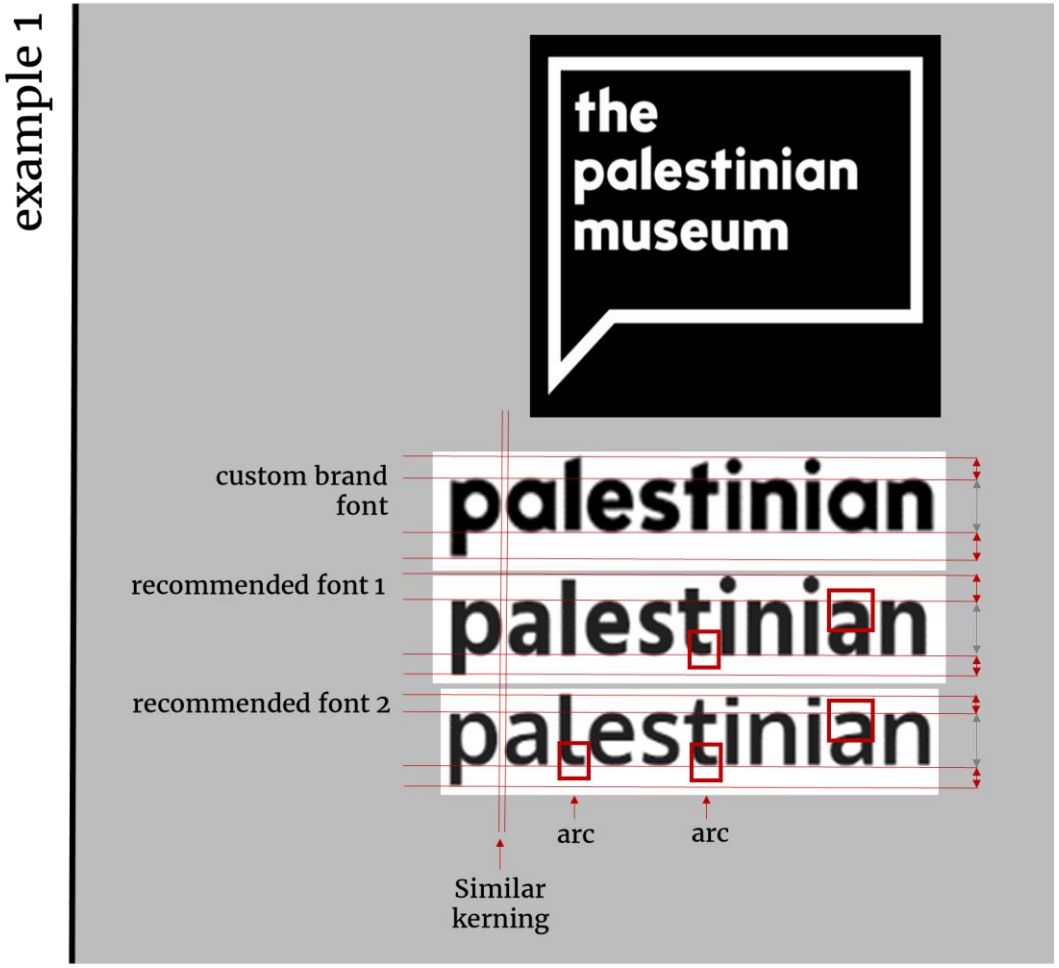
Here is your recommended font!

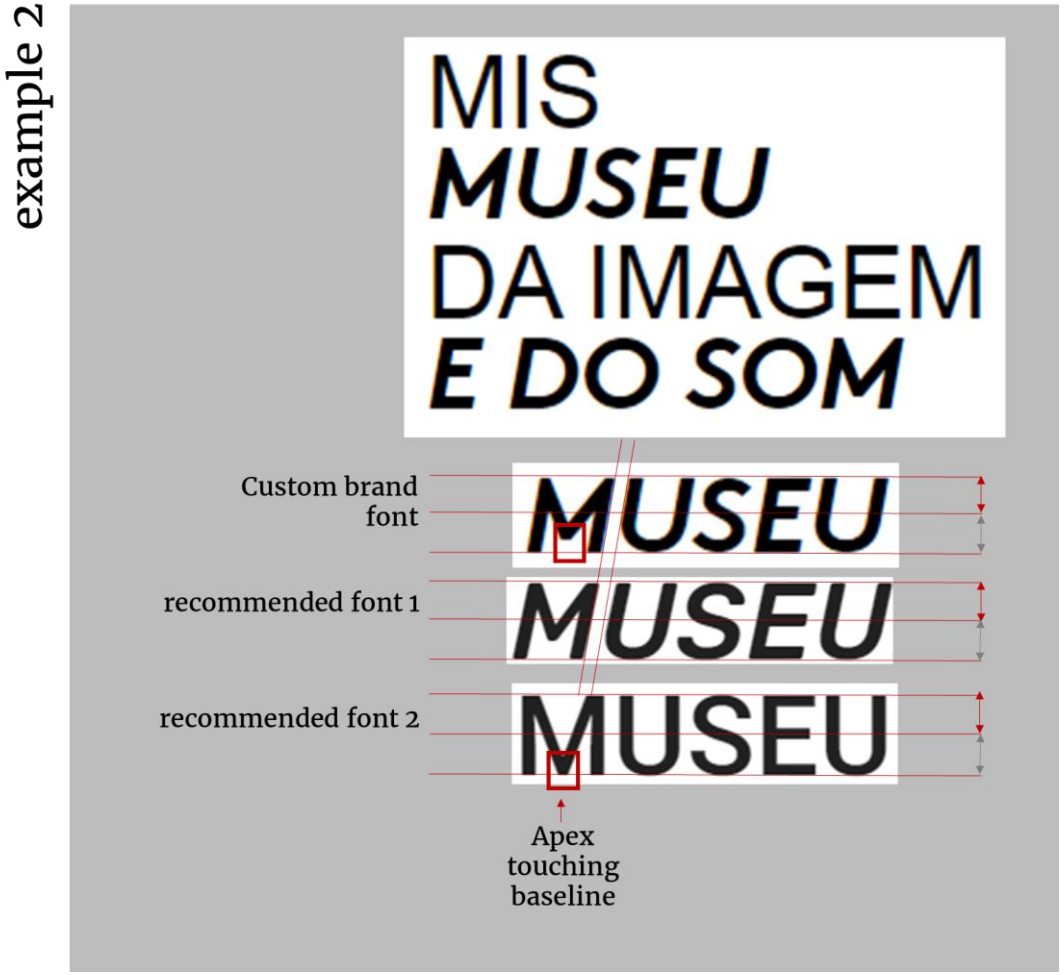
Font: **Oxygen**

the palestinian museum
the palestinian museum
the palestinian museum

Example

Our test case was taken from the
International Society of Typographic Designers Certificate of Excellence Awardees





Demo

face of fonts 字之魂

the recommendation:

enter mission here:

enter ideal font here:

Get font

analysis

main font

Demo

face of fonts 字之魂

enter mission here:

e with reliable technology

enter ideal font here:

Get font

the recommendation:

PT Sans regular

cosine similarity:

analysis

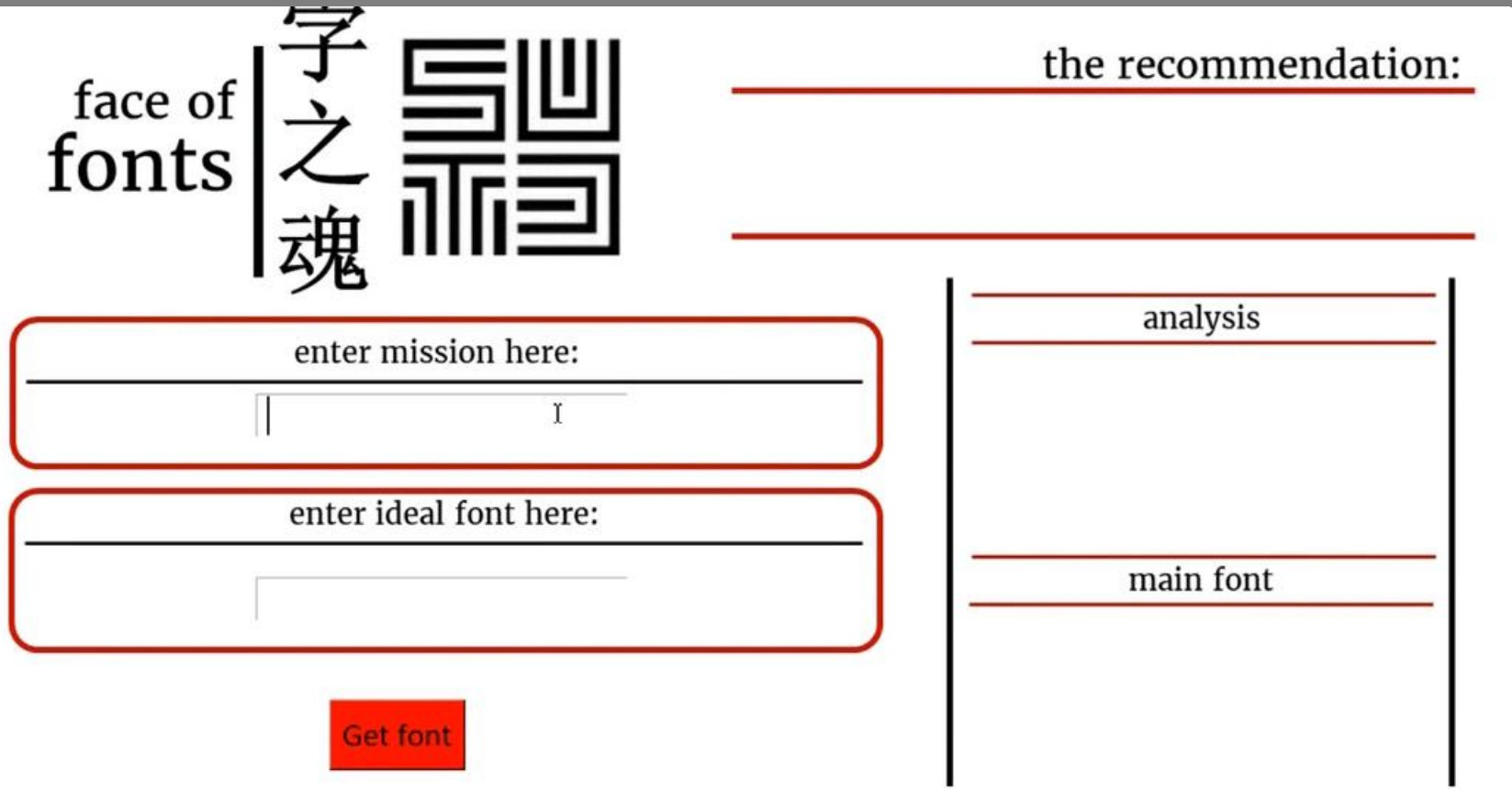
Helvetica Nueue	0.095361	Helvetica	0.0775914
Arial	0.007729	Myriad	0.467078
Futura	0.003081	Gotham	0.292116
Frutiger	0.010451	Gill Sans	0.0395129
Garamond	0.00707802		

main font

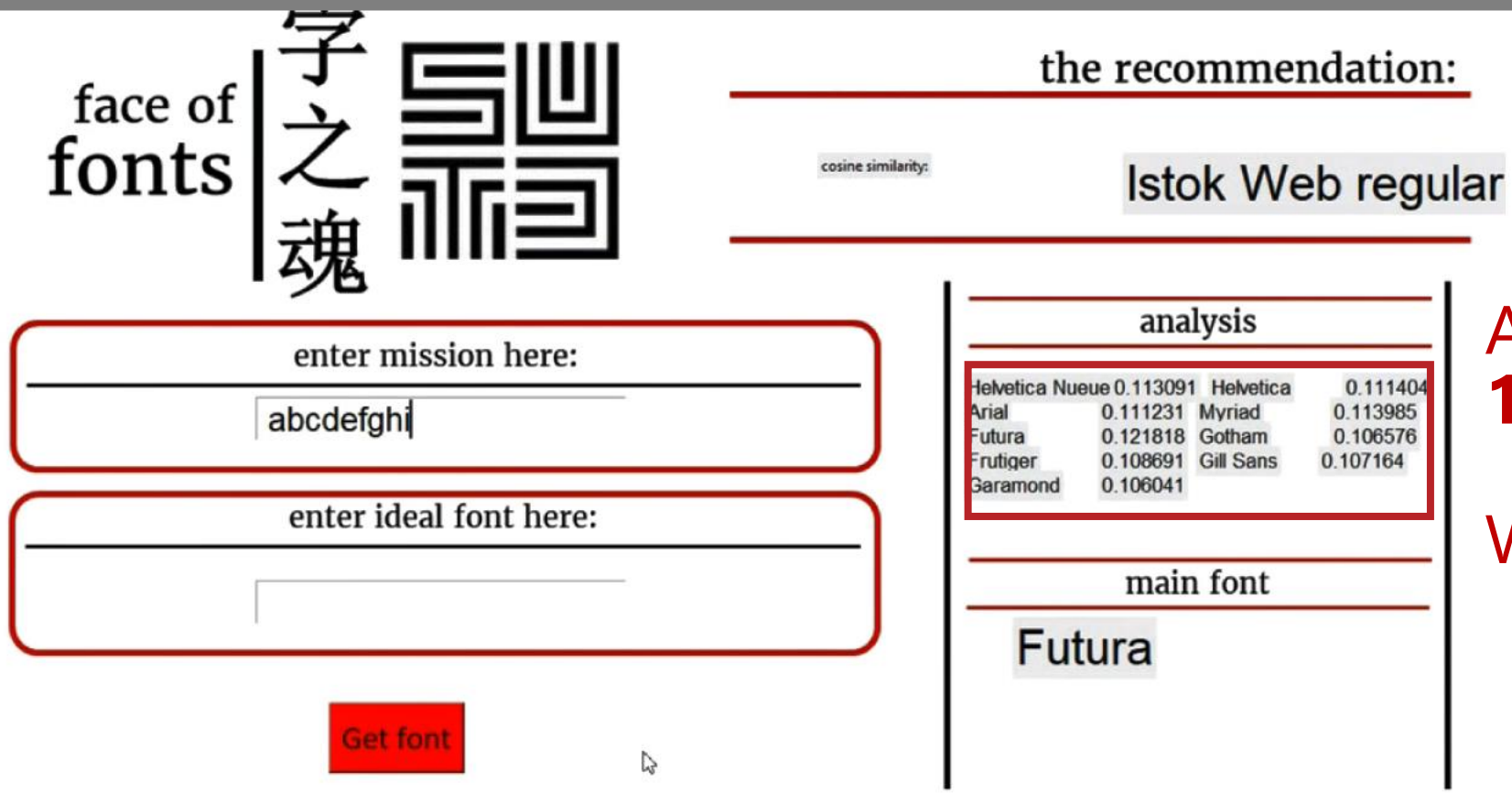
Myriad

Laseg
dhum
Hloiv

Demo


face of
fonts
字之魂
the recommendation:
analysis
enter mission here:
enter ideal font here:
main font
Get font

Demo



All the probabilities are **11%**

Which is random chance

Summary

input
Paragraph describing corporation or project

Data-aided recommendation by CNN

intermediary
Probability makeup of each font

composition
Vector composition of vectors using probability estimates as weights

output
Obtain top 3 fonts that are closest by Euclidean or Cosine distance

Conclusion

1. First time working on semantic of typography, hypothesis holds true
2. System for intelligence augmentation using data-driven contextual analysis
3. Build larger and diverse dataset across languages and domains
4. Refine font recommender with designers according to their needs