# User Recommenders, For Scottsdale AZ

Four Recommender from Yelp's Open Dataset

March 2020

David Yu

https://bit.ly/3200bVk

# **Data Source: Yelp Open Dataset**











6,685,900 reviews

192,609 businesses

200,000 pictures

10 metropolitan areas

1,223,094 tips by 1,637,138 users Over 1.2 million business attributes like hours, parking, availability, and ambience Aggregated check-ins over time for each of the 192,609 businesses

Source: yelp.com/dataset

# **Motivation: Why Scottsdale?**



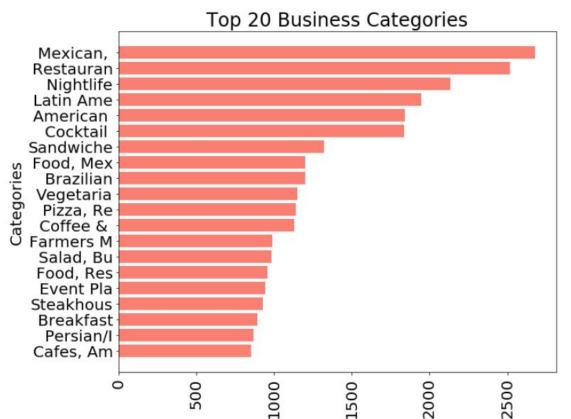




Source Forbes: https://bit.ly/2vAIXON

# **Businesses in Scottsdale**





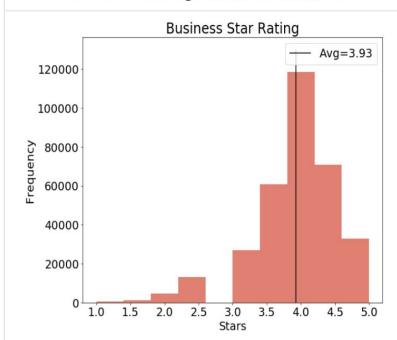
7,081

Total
Businesses in dataset
(Span multiple categories)

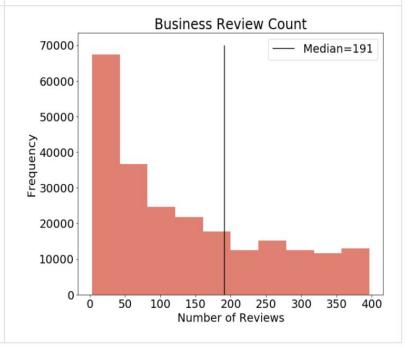
# **Businesses Ratings in Scottsdale**



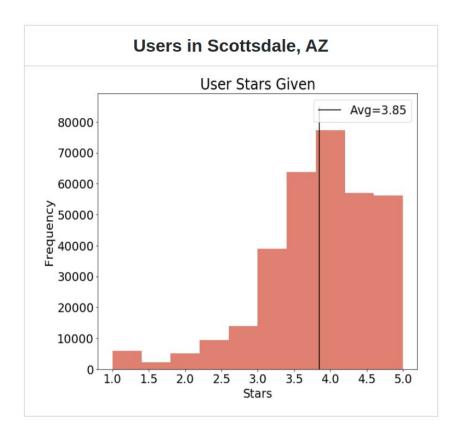




### **Business Review Count in Scottsdale, AZ**



# **Users in Scottsdale**





### **KEYNOTE:**

Average *user rating* will be used as the *metric* for subsequent recommenders

### **RATIONALE:**

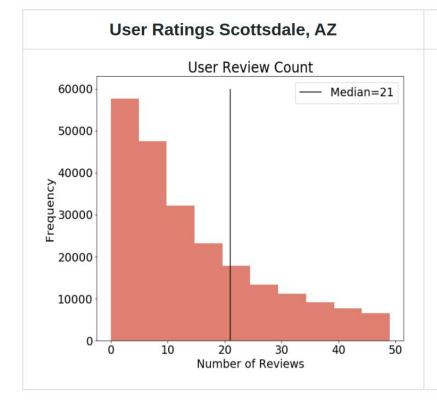
Business ratings can be used but some users patronize *poorly rated* businesses

140,940

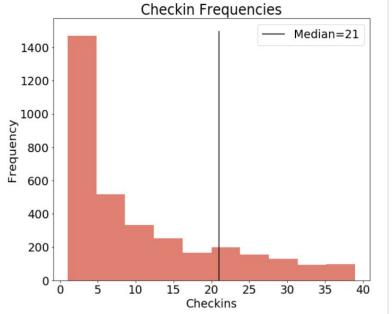
**Unique User ID's** 

# **Users in Scottsdale**



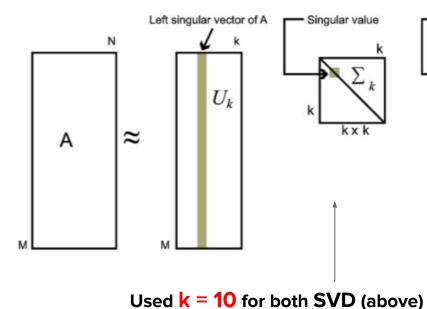


## User Check-In Counts Scottsdale, AZ



# **Math: SVD and NMF**





and NMF (right)

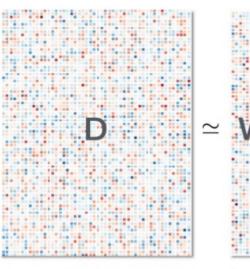
Right singular vector of A

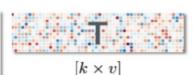
N

k

V<sub>k</sub>

T





Matrix Factorization

**SVD** Image Source: <a href="https://bit.ly/3ctCYMj">https://bit.ly/3ctCYMj</a> **NMF** Image Source: <a href="https://bit.ly/3coElgd">https://bit.ly/3coElgd</a>

 $[n \times v]$ 

 $[n \times k]$ 

# **SVD** and **NMF** Top 10 Predictions



SVD Recommendations in	NMF Recommendations in	Patronized Establishments in
Scottsdale, AZ	Scottsdale, AZ	Scottsdale
('Kristin', ['True Food Kitchen', 'Cafe Monarch', 'The Herb Box', 'Sweet Republic', 'Talking Stick Resort', 'Roka Akor - Scottsdale', 'Sprinkles Cupcakes', "Portillo's Hot Dogs", 'iPic Theaters', 'Studio Movie Grill'])	('Kristin', ['True Food Kitchen', 'Hopdoddy Burger Bar', 'The Thumb BBQ', 'SOL Mexican Cocina', 'Butters Pancakes & Café', "Coconut's Fish Cafe", 'Salty Sow', 'North Italia', "Portillo's Hot Dogs", 'Sweet Republic'])	3360 Butters Pancakes & Cafe 312378 Wildflowe 600627 Flower Child 624904 Viet Kitche 744479 Wildflowe 746301 Original Breakfast Hous 771054 Andreoli Italian Groce 838710 Barrio Quee 1147492 Rocket Burger & Sub 1155989 Rustic Cafe

# **SVD** and **NMF** Validations



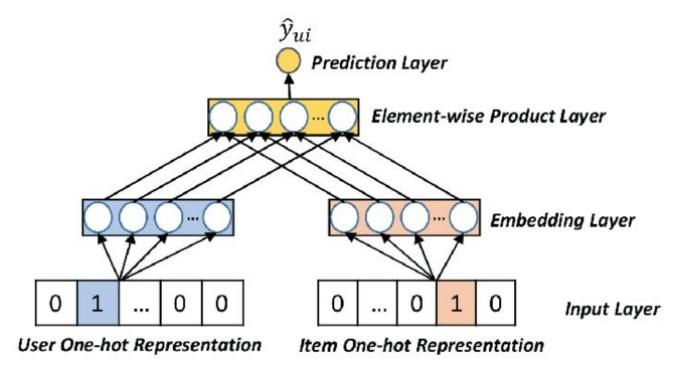
7 CV Validations with Surprise k=10			
	test_rmse	test_mae	
Algorithm			
SVD	0.691716	0.478392	
NMF	0.703249	0.536901	
BaselineOnly	0.715314	0.506353	
NormalPredictor	1.101105	0.865834	



**SVD** and **NMF** models are not off by much

# **Embedded Neural Network Model**





Embed Image Source: https://bit.ly/2VEbDkF

```
Embedded Neural Network Results
array(['Butters Pancakes & Café', 'Wildflower', 'Flower Child',
      'Andreoli Italian Grocer', 'Barrio Queen',
      'Grassroots Kitchen & Tap', 'La Fonda Del Sol', 'Diego Pops',
     "Jason's Deli", 'Zipps Sports Grill', 'Arrivederci', "Corleone's",
```

'Jalapeno Inferno', 'Mozaik Skin & Body', 'NYPD Pizza', 'NYPD Pizza', 'Naked BBQ', "Casella's Italian Delicatessen",

"Casella's Italian Delicatessen", "D'Lite Healthy On The Go", 'Dairy Queen', 'Asian Island', "Raising Cane's",

'Wow Wow Lemonade Scottsdale', "Coconut's Fish Cafe",

'Paradise Hills Dentistry', 'Paradise Hills Dentistry',

'Connected Medical Center'], dtype=object)

array(['John Gluch - Platinum Living Realty', 'Riddle Escape Room', 'Wave House Cleaning Service', 'Nu Nails', 'Sasaki Sushi & Bar',

'Just Smiles', 'Plush Nail Bar', 'Van Buick GMC',

'Chapman Hyundai Scottsdale', 'Olga Brow Studio', 'Pangaea Dinosaur Grill', 'Better Than New Furniture Repair',

'Vidogi Salon', "Children's Dental Specialty Group",

'The Glam Nurse', 'Authentic Thai Massage',

'Pink Petals Nails Spa', 'National Laser Institute',

'The Hill Group',

'Affiliated Pediatric Dentistry & Orthodontics - Grayhawk',

'Four8Zero', 'We Do Men - A Male Concept Spa', 'Rolfs Salon At Gainey Village',

'Whiten Up Today Laser Teeth Whitening', 'Fabulous Cuts',

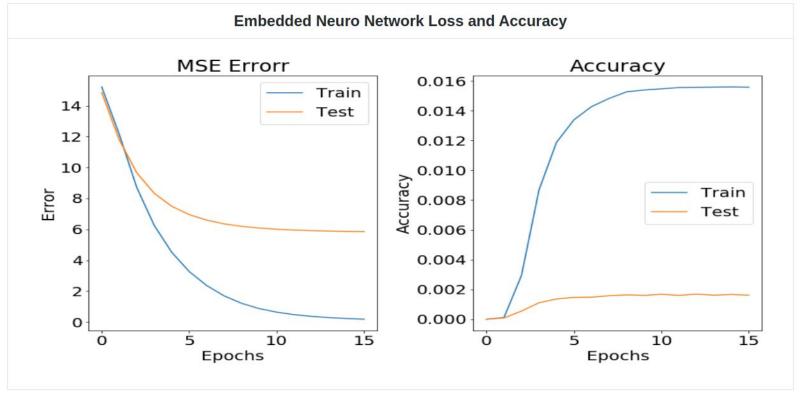
'DeRosa Medical', 'Taj Mahal', 'R&R Law Group'], dtype=object)

Kristin's **Patronage** 

Kristin's Recommendations

# **Embedded Neural Network Results**

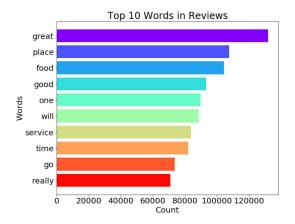


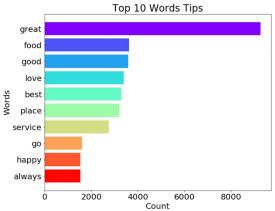


# **Recap and Next Steps**



- 1. **SVD** and **NMF** models made better historical recommendations
- Neural Network model performance can be improved with more data or prior training
- Investigation into model accuracy for top n recommendations instead of just one





# **Questions**:





