



Problem

- Given a product's ingredient list, can we more or less know its category?
 - Are eye creams and face creams substantially different in terms of their formula or they are created mostly for a marketing purpose.
 - Classification of product category with ingredient information.
- How important is product formula in determining a beauty product's price?
 - Is a product expensive because it has sophisticated ingredients, or because it has attractive packaging, or simply because it is from a high-end brand.
 - Regression on price with ingredient or non-ingredient features.

Data scraped from reviewing website [beautypedia.com](https://www.beautypedia.com)

Product Table

	product_links	product_names	product_category	brand	ingredient	size	price	claims	image_pat
0	https://www.beautypedia.com/products/clear-complexion-spot-treatment	Clear Complexion Spot Treatment	Acne & Blemish Treatment	Merle Norman	Active: Salicylic Acid (2%), Other: Alcohol D...	0.50 fl. oz.	20.00	Fast-drying, on-the-spot solution with Salicyl...	images/skin_care/Clear-Complexion-Spot-Treatme.
1	https://www.beautypedia.com/products/acne-solutions-emergency-gel-lotion	Acne Solutions Emergency Gel Lotion	Acne & Blemish Treatment	Clinique	Active: Benzoyl Peroxide (5%), Other: Water P...	0.50 fl. oz.	17.00	Clinique's medicated troubleshooter with benzo...	images/skin_care/Acne-Solutions-Emergency-Gel-
2	https://www.beautypedia.com/products/resist-daily-pore-refining-solution-2-bha	RESIST Daily Pore-Refining Solution 2% BHA	Acne & Blemish Treatment	Paula's Choice Skincare	Water (Aqua), Dipropylene Glycol, Salicylic ...	3.00 fl. oz.	33.00	This daily-use liquid exfoliant is ideal for t...	images/skin_care/RESIST-Daily-Pore-Refining-So.
3	https://www.beautypedia.com/products/max-complexion-correction-pads	Max Complexion Correction Pads	Acne & Blemish Treatment	Peter Thomas Roth	Active: Salicylic Acid (2%), Other: Alcohol D...	60.00	40.00	Regulates healthy cell turnover and delivers p...	images/skin_care/Max-Complexion-Correction-Pad.
4	https://www.beautypedia.com/products/naturals-acne-spot-treatment	Naturals Acne Spot Treatment	Acne & Blemish Treatment	Neutrogena	Active Ingredient: Salicylic Acid 1%. Inactive...	0.75 fl. oz.	8.49	Clinically proven acne medicine treats and hel...	images/skin_care/Naturals-Acne-Spot-Treatment_.

Ingredient Table

	name	category	rating	rating_num
0	1, 2-Hexanediol	[Preservatives]	Good	2
1	10-Hydroxydecanoic Acid	[Emollients]	Good	2
2	4-T-butylcyclohexanol	[Emollients, Skin-Soothing]	Good	2
3	Acacia farnesiana extract	[Plant Extracts, Fragrance: Synthetic and Frag...	Poor	0
4	acacia senegal gum	[Texture Enhancer, Plant Extracts, Skin-Soothing]	Good	2

- 8000+ products
- 1750 ingredients

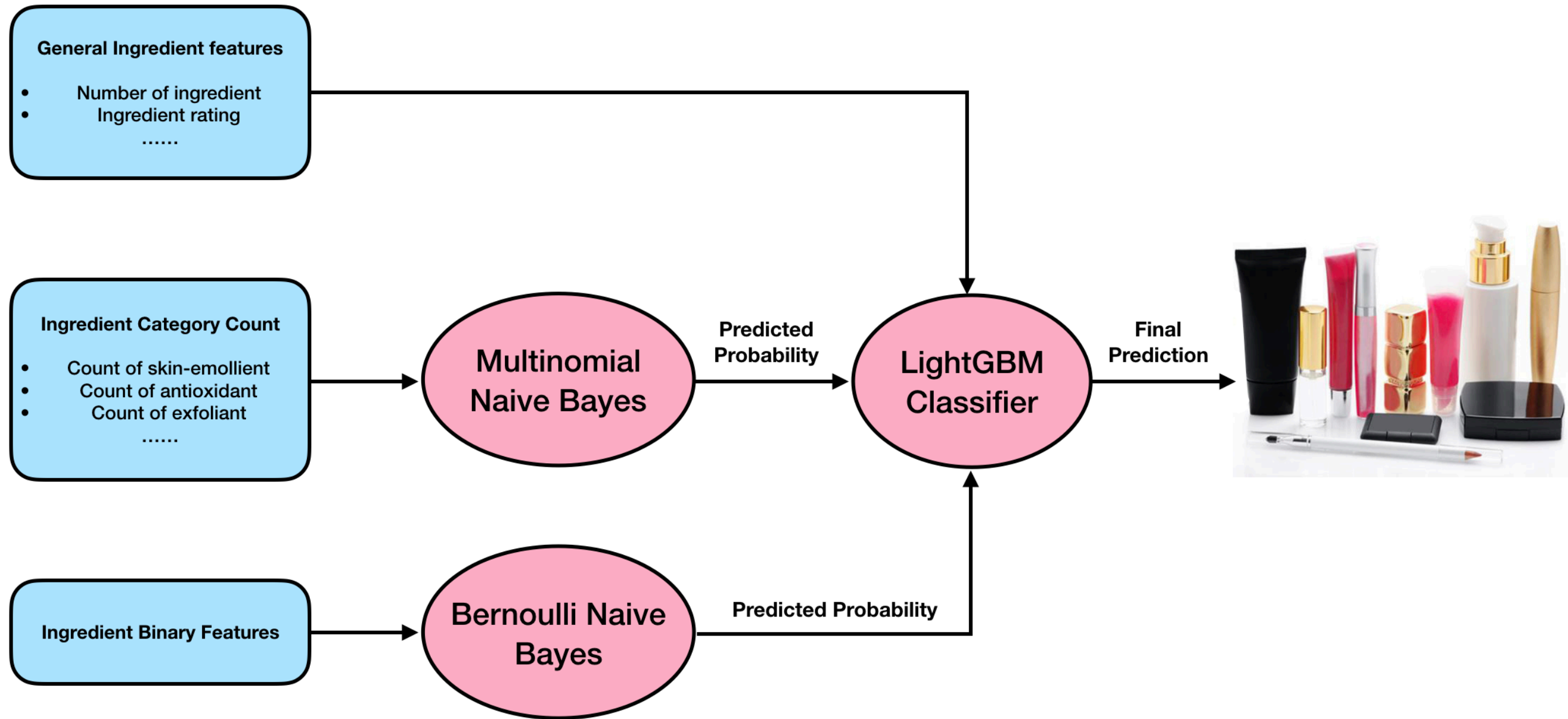
Ingredient Matching

- Different companies may list the same ingredient in different ways. The most common ingredient water, can appear as “water”, “water (aqua)”, “Water/Aqua/Eau”, “purified water”...
- We want to make use of the information in ingredient table.
- Find best matching ingredient that rating and category are available using use python function SequenceMatcher, which evaluate a similarity metric between two strings.
- After matching, create multiple features such as ingredient category counts, average ingredient rating...

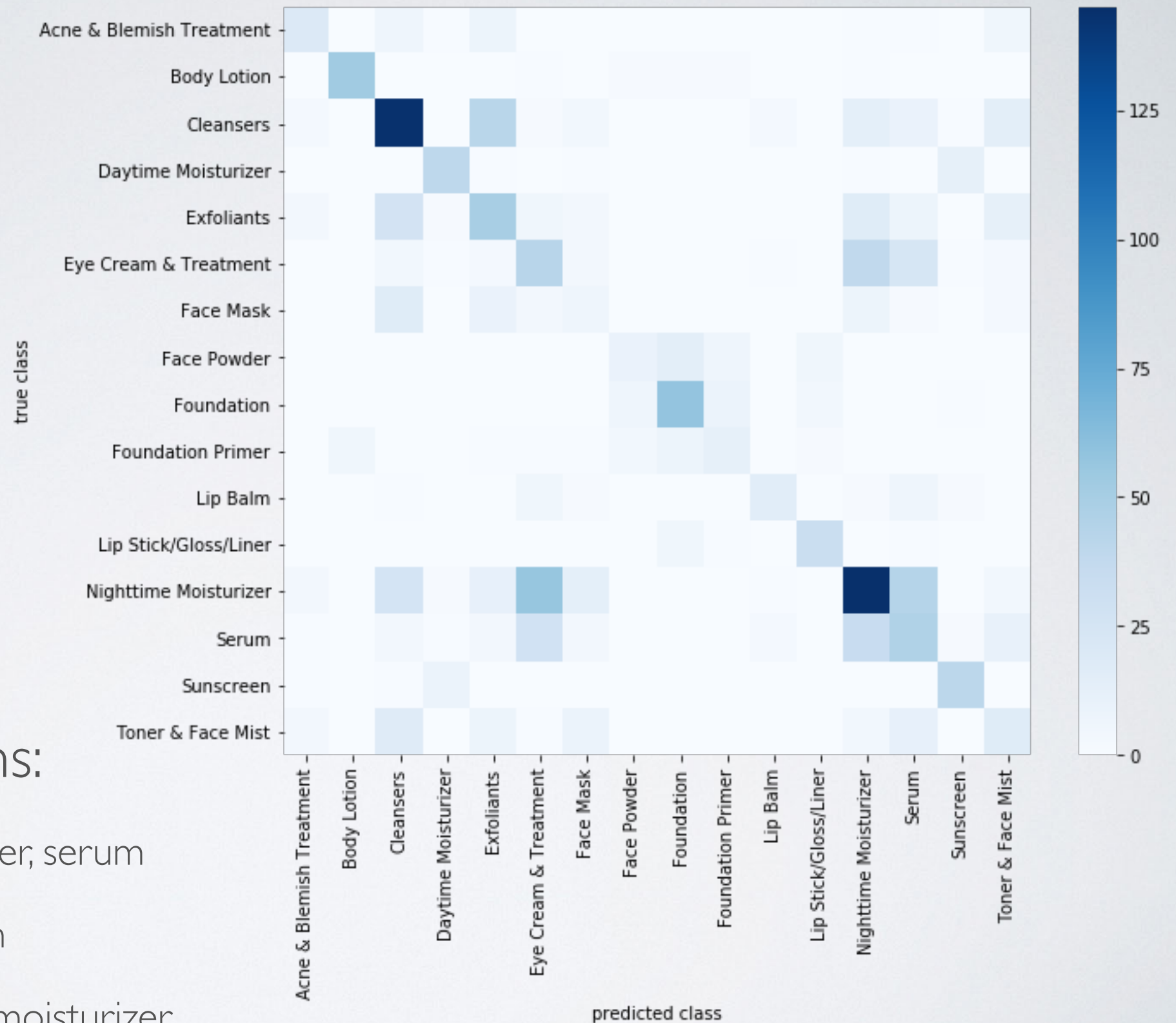
Example of Ingredient Matching Results

		matching	rating	category
Coleus Forskohlii Root Oil	Carthamus tinctorius oil		3.0	[Plant Extracts, Emollients]
Sodium Ascorbyl Phosphate	sodium ascorbyl phosphate		3.0	[Skin-Soothing, Vitamins, Antioxidants]
Lavandula Hybrida (Lavender) Oil	Lavandula angustifolia		0.0	[Plant Extracts, Sensitizing, Fragrance: Synth...
Cocos Nucifera (Coconut) Water	cocos nucifera (coconut) fruit extract		3.0	[Skin-Softening]
Lactic Acid	acetic acid		0.0	[Sensitizing]
Glycosaminoglycans	glycosaminoglycans		3.0	[Skin-Replenishing, Skin-Restoring]
Sodium Cetearyl Sulfate	sodium cetearyl sulfate		2.0	[Cleansing Agents]
Cetareth-20	cetareth-20		2.0	[Texture Enhancer]
Camellia Sinensis (Green Tea) Leaf Extract	Scutellaria baicalensis extract		2.0	[Skin-Soothing, Plant Extracts, Antioxidants]
Niacinamide	niacinamide		3.0	[Skin-Soothing, Skin-Restoring, Antioxidants, ...]
Magnesium Ascorbyl Phosphate	magnesium ascorbyl phosphate		3.0	[Antioxidants, Vitamins]
Glycerin	glycerin		3.0	[Skin-Replenishing, Skin-Restoring]
Fragrance (Parfum).	fragrance		0.0	[Fragrance: Synthetic and Fragrant Plant Extra...
1,2-Hexanediol	1, 2-Hexanediol		2.0	[Preservatives]
Peanut Oil	peanut oil		2.0	[Plant Extracts, Emollients]
Isopropyl Myristate	isopropyl myristate		2.0	[Texture Enhancer, Emollients]
Stearyl Glycyrrhetinate	stearyl glycyrrhetinate		3.0	[Skin-Soothing, Plant Extracts]
Sodium Borate	sodium borate		0.0	[Sensitizing]
Carthamus Tinctorius Oil/Safflower Seed Oil	Carthamus tinctorius oil		3.0	[Plant Extracts, Emollients]
Melaleuca Alternifolia (Tea Tree) Oil	Melaleuca alternifolia		2.0	[Plant Extracts, Antioxidants]
Hydroxypropyl Cyclodextrin	hydroxypropyl cyclodextrin		2.0	[Miscellaneous]
Olea Europaea (Olive) Fruit Extract	Olea europaea fruit oil		2.0	[Antioxidants, Emollients, Plant Extracts, Ski...
Palmitoyl Tripeptide-38.	Palmitoyl Tripeptide-38		3.0	[Skin-Restoring]
Sodium Hydroxide	sodium hydroxide		1.0	[Cleansing Agents, Sensitizing]
Zea Mays (Corn Germ) Oil	wheat germ oil		2.0	[Plant Extracts, Emollients]
Dimethiconol	dimethiconol		2.0	[Silicones, Emollients]

Product Category Classification



Confusion Matrix



Typical confusions:

- Eye cream, nighttime moisturizer, serum
- Daytime moisturizer, sunscreen
- Face mask: cleanser, nighttime moisturizer

- **Clinique Moisture Surge™ Overnight Mask**

- predicted as Eye Cream & Treatment and Nighttime Moisturizer

ingredient: Water, Glycerin, Cetyl Alcohol, Dimethicone, Glyceryl Polymethacrylate, Butyrospermum Parkii (Shea Butter), Cetyl Ethylhexanoate, PEG-8, Glycereth-28, Sucrose, Sorbitan Stearate, PEG-100 Stearate, Trehalose, Mangifera Indica (Mango) Seed Butter, Hypnea Musciformis (Algae) Extract, Gelidium Acerosa (Algae) Extract, Olea Europaea (Olive) Fruit Extract, Triticum Vulgare (Wheat Bran) Extract, Cladosiphon Okamuranus Extract, Astrocaryum Murumuru Seed Butter, Cetearyl Alcohol, Aloe Barbadensis Leaf Water, PEG-75, Caffeine, Pantethine, Sorbitol, Butylene Glycol, Oryzanol, Bisabolol, Panthenol, Phytosterols, Tocopheryl Acetate, Caprylyl Glycol, Sodium Hyaluronate, Hexylene Glycol, Carbomer, Potassium Hydroxide, Dextrin, Disodium EDTA, Phenoxyethanol, Red 4, Yellow 5



This mask has moisturizing ingredients such as glycerin, shea butter, and a number of antioxidant. This makes it similar to a moisturizer and eye cream.

- **Clinique Pep-Start Double Bubble Purifying Mask**

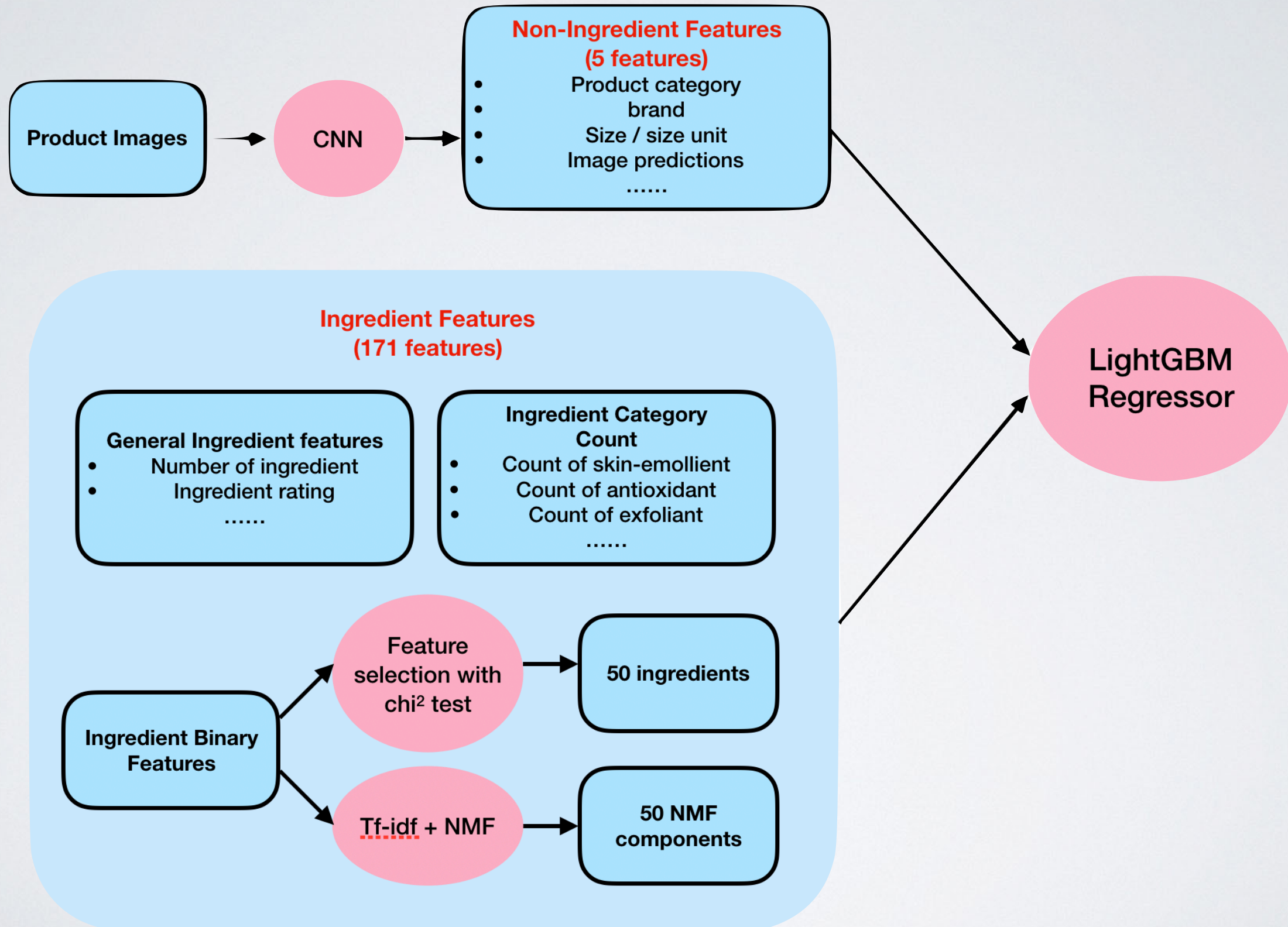
- predicted as Cleansers.

ingredient: Water, Disiloxane, Cocamidopropyl Betaine, Sodium Cocoyl Isethionate, Decyl Glucoside, Glycerin, Pentylene Glycol, Acrylates Copolymer, Sodium Chloride, Morus Nigra (Mulberry) Root Extract, Scutellaria Baicalensis Root Extract, Vitis Vinifera (Grape) Fruit Extract, Palmitoyl Tetrapeptide-7, Palmitoyl Tripeptide-1, Acetyl Glucosamine, PEG-6 Caprylic/Capric Glycerides, Ethylhexylglycerin, Butylene Glycol, Polysorbate 20, PEG-150 Pentaerythrityl Tetrastearate, Potassium Hydroxide, Coconut Acid (Coconut Derived), Acrylates/C10-30 Alkyl Acrylate Crosspolymer, Carbomer, Xanthan Gum, Disodium EDTA, Phenoxyethanol, Red 33



The cocamidopropyl betaine and sodium cocoyl isethionate in this product are surfactants, which makes it easy to be confused with cleansers.

Price Regression



Prediction Results

error\features	non-ingredient features only	ingredient features only	all features
RMSE(\$) (train cv)	22.280	27.445	21.393
RMSE(\$) (test)	18.629	24.803	17.100
MAE(\$) (train cv)	12.322	17.130	11.692
MAE(\$) (test)	10.875	16.298	10.257
explained variance (train cv)	0.610	0.408	0.640
explained variance (test)	0.683	0.438	0.732

Non-ingredient features are more powerful but ingredient features are not useless (good to know!)

Thank you for your time!

Project Repo: https://github.com/NoxMoon/inside_beauty