Inspiring healthy habits: data science at WW

Carl Anderson

carl.anderson@ww.com
@leapingllamas
Outline

- Intro to WW: purpose, program, type and scale and data
- Behavioral Nudges
- WW Data Products
- Primrose: how we develop and deploy ML models
- Q&A
About Me

Data Strategy

Data Science

Business Intelligence Engineering

Wellness that Works.
40% of adults on the planet are already overweight or obese and more are joining their ranks every day.

While the world’s bulging waistlines are driven by economic success—wealthy populations eat more—obesity’s estimated cost of $2 trillion a year worldwide is now almost as much of a financial burden as smoking.

Top 10 countries ranked by percentage of adult population that is obese in 2015 or nearest year:

1. US - 41.0%
2. Mexico - 37.5%
3. NZ - 32.0%
4. Hungary - 31.5%
5. Chile - 30.7%
6. Australia - 27.4%
7. UK - 26.8%
8. Finland - 25.8%
9. Canada - 25.5%
10. Germany - 23.9%

HEALTH PARADOX #1

We spend more time and money than ever before on wellness, but we’ve never been more unhealthy.
HEALTH PARADOX #2

Despite all the advances in science and food production, eating healthy has never felt more complicated.
Confusing headlines have left many in a fog and unsure what to do.

<table>
<thead>
<tr>
<th>Pro</th>
<th>Con</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 Reasons You Need to Count Calories</td>
<td>5 Reasons To Never Count Another Calorie</td>
</tr>
<tr>
<td>Why Sugar isn’t the Bad Guy</td>
<td>10 Disturbing Reasons Why Sugar is Bad for You</td>
</tr>
<tr>
<td>Why Gluten is Evil</td>
<td>JUNK SCIENCE: Gluten is Not Bad For You</td>
</tr>
<tr>
<td>Grain and Autoimmune Disease</td>
<td></td>
</tr>
<tr>
<td>THE SECRET IS OUT: THE REASON RED MEAT IS BAD FOR OUR HEALTH</td>
<td>RED MEAT: It Does a Body Good!</td>
</tr>
<tr>
<td>THE ORGANIC FOOD LIE</td>
<td>New study finds organic foods are healthier than conventionally grown foods</td>
</tr>
</tbody>
</table>
People want inspiration, not just information.

People want healthy habits that fit their lives.

Today, healthy is the new skinny.

It’s not about lifestyle, it’s about livability.

Community is essential.

People crave purpose.
OUR PURPOSE

We inspire healthy habits for real life.*

* For people, families, communities, the world—for everyone.
WW: Wellness Ecosystem

- Communities
- Nutrition
- Motivation
- Activity
- Mindset
- Stress
- Sleep

Wellness that Works.
TENET #1
We help you build powerful habits, rooted in science.

TENET #2
We know you and meet you where you are.

TENET #3
We enable you to find and form inspired communities.
SmartPoints is about health, not just calories.

All calories are NOT created equal.

- Greek yogurt & fruit and peanut parfait (1 SmartPoints value)
- Scrambled eggs, Canadian bacon, avocado, tomato and English Muffin (6 SmartPoints value)
- Pancakes with chocolate chips and maple syrup (12 SmartPoints value)

Wellness that Works.
Nutritional science to make healthy eating simpler

SmartPoints nudges you towards a healthy eating pattern with more fruits, vegetables and lean protein, and less sugar and saturated fat.

- **Calories** establish the baseline.
- **Sugar** and **Saturated Fat** increase the SmartPoints value.
- **Protein** lowers the SmartPoints value.
- Foods that form the foundation of a healthy eating pattern have SmartPoints value of zero.
ZeroPoint™ Foods

ZeroPoint foods form the foundation of a healthy eating pattern and have a low risk of overeating.

They don’t have to be weighed, measured, or tracked.

- Vegetables
- Fruits
- Skinless chicken breast
- Skinless turkey breast
- Non-fat plain yogurt
- Eggs
- Beans, peas, lentils, tofu and corn
- Fish & seafood
Everything is on the menu
WW Studio

30,000 meetings per week globally
FREE SHIPPING on your entire purchase. Offer details.

Mix and match a perfect batch.
Mini bars are 2 for $12* with code MINISALE

Shop now

*Offer valid on any two individual boxes of Dark Chocolate Raspberry, Chocolate Caramel, Chocolate Pretzel, Mint, Cookie Bar or So Good! Snack Mini Bars. Use promo code MINISALE. Offer available while supplies last. You must enter promo code exactly as shown at checkout to receive discount. $12 is the total value for two boxes. Offer valid through September 29, 2023. Offer may expire without notice due to error, fraud or other unforeseen circumstances.

NEW ARRIVALS

Weight Watchers Magazine September/October Issue $4.95
Buttermilk Protein Pancake $4.50
Best of WW Mini Cookbook Bundle $11.95
Snickerdoodle Baked Protein Mini Bar $5.95
Triple Chocolate Baked Protein Mini Bar $5.95
Butter Popcorn - Pack of 6 $7.95
Activity & Mindset

WW & Headspace have teamed up to offer short mindfulness techniques you can do any time, anywhere.

Changing Perspectives
Taking time to be mindful can make a big difference.
Voice & AI

Google Assistant

Alexa

Wellness that Works.
Get ready to feel your strongest
This 30-day plank challenge gives you a new goal (and a toned core) to work toward. Are you in?

FUEL UP AFTER WORKING OUT
1. Recipes for people who love peanut butter
2. The best protein bars to eat on WW
3. Protein-packed meals and snacks

BEAT BACK PAIN
This multi-tasking move helps to ease an achy back, fix posture, boost balance, and more.
Proven digital health solution for teens and children

A new app from @ww.us aims to inspire healthier habits among kids, teens, and their families. @mindygrossman, CEO of WW, hopes the app provides families with "a more comprehensive way to support their efforts" toward the goal of living a healthier life.

"I feel more confident, healthier, and more comfortable in my skin."
In Q2,
Social network:
● 2 million posts
● 14 million comments
● 70 million likes

1 million members tracked a physical activity
Our geographic footprint

- Canada
- USA
- Brazil
- France
- UK
- Benelux
- Germany
- Sweden
- Switzerland
- Australia
- New Zealand
Member Journey (Illustrative)

- Visitor
- Sign up
- On board
- Track food
- Read article
- Do activity
- Buy snacks on e-comm
- Go on cruise
- Personal coach
- Post on Connect
- Purchase snacks
- Attend meeting
- Meditate
- Try meal kit
- Call Customer service
- Redeem win
- Make recipe
- Use Alexa assistant

Almost none of this is personalized!
Big data:
- Food
- Activity
- Exercises
- Challenges
- Social network
- Workshops
- Personal Coaches
- CRM
- Fulfillment
- Meal kits
- Supermarket foods
- E-commerce
- Cruises

...for 56 years

Wellness that Works.
Scale of Data

- **Nackers et al. (2013)** showed that fast ($\geq 0.68$ kg/week) weight loss in the first month predicts higher weight loss success at 6 months than slow (<0.23 kg/week) or moderate initial weight loss.

- Sample size: 298

- We checked our weigh in data to compare these results to what we observe in our member base.

Scale of Data

- Members considered:
  1) started and ended their membership between Apr 2017 and Apr 2019
  2) were members for at least 6 months
  3) weighed in in their first week, fourth week and sixth month
  4) were obese at the beginning of their membership (BMI > 30)
- For all analyses (mostly) unfiltered self report data was used.

Sample size: 211,000 members!
7kg median weight loss after 6 months

- 211k members
- 54% digital (sampling bias)
- Mostly female
- Median birth year 1971
- Median start BMI: 35
- Median BMI @ 6 months: 32.5
Significant effect of initial weight loss rate on weight loss success

- Weight loss defined as 5% or 10% of initial start weight lost
- Initial weight loss speed:
  - Fast (≥0.68 kg/week): 116,107
  - Moderate (<0.68 & >0.23 kg/week): 61,204
  - Slow (<0.23 kg/week): 34,107

Proportion of members who lost 5% or 10% initial body weight at 6 mo

*all differences statistically significant
Nudges & behavioral change
Healthy Habits

Habit:

“Automatic behaviors triggered by situational cues”

“Habit-forming companies link their service to the users’ daily routines”
“Any addition to or modification of the environment that influenced consumers in a predictable way, without changing economic incentive”

Altering environment by changing presentation of options — called **choice architecture**
“Any aspect of the choice architecture that alters people’s behavior in a predictable way (1) without forbidding any options or (2) significantly changing their economic incentives. Putting fruit at eye level counts as a nudge; banning junk food does not.”
Nudging & Choice Architecture

iNcentives
Understand mappings
Defaults
Give feedback
Expect error
Structure complex choices

This is Thaler & Sunstein's framework. Instead, I will use Blumenthal-Barby & Burroughs
## Nudging & Choice Architecture

<table>
<thead>
<tr>
<th>Category</th>
<th>Explanation</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Priming</strong></td>
<td>Subconscious; physical, verbal, sensational</td>
<td>Place unhealthy options out of sight or farther away in cafe</td>
</tr>
<tr>
<td><strong>Salience</strong></td>
<td>Informational; attention grabbing; emotional</td>
<td>Calorie label; graphic image on cigarette cartons; recommenders</td>
</tr>
<tr>
<td><strong>Default</strong></td>
<td>Pre-set default choice; good option for do-nothing behavior</td>
<td>Automatic opt-in, have to explicitly change or opt-out: benefits, organ donation</td>
</tr>
<tr>
<td><strong>Incentive / Gamify</strong></td>
<td>Reward or punish for behavior; recognition</td>
<td>Badge; coupon; $$?</td>
</tr>
<tr>
<td><strong>Commitment / Ego</strong></td>
<td>Get someone to make a commitment; leverage ego, pride</td>
<td>Sign up for 5k; invest (pay for membership); share with friends</td>
</tr>
<tr>
<td><strong>Norms / Messenger</strong></td>
<td>Use other to establish a norm and for consumer to compare themselves</td>
<td>80% of (other) people are organ donors; most people wear seatbelts</td>
</tr>
</tbody>
</table>

Everything is on the menu

Members are empowered with lots of choices.
We are not dictating diet, exercise regime etc. Hence, these are nudges.
Priming
Visibility, accessibility, available

Further prime by reducing friction for key actions such as tracking
We plan to leverage notification nudges, similar to these, to help keep people on track.
Priming

Visibility, accessibility, available

Connect social network: very highly supportive, lots and information from staff and fellow members. Priming and saliency.
Points on very large number of foods

Clear “mappings” make it easier to make good choices:
  ○ instead of calorie counting, 300 cal (or is it kJ, kcal) → 5 points
Saliency
meaningful, relevant info

- Clear budget
- Clear progress
Saliency
meaningful, relevant info

Kurbo traffic light system

UK nutrition labels

Source: Food Standards Agency
Saliency
meaningful, relevant info

Activities are also pointed: FitPoints
Saliency
meaningful, relevant info

Tracking and wearables
Santorio, Santorio
1561-1636
Tracker

- Wake up early
- Morning routine
- Meditation
- Workout
- Bread only
- Hydrate
- Budget review
- Blog post
- TT video
- Instagram
- "One day"
- Bed by 11pm

Mood:
- Happy
- Excited
- Stressed
- Tired
- Productive
- Grateful

# Gratitude
Wellness that Works.
Defaults

- Defaults should be good, fair, equitable...
- Easily changed
Incentives / Gamify
Incentives / Gamify

Blue dot: daily, weekly, monthly
Week of Aug 26 - Sep 1

11
FitPoints Remaining
0
Daily Remaining
26
Daily Used

**Breakfast total 2**
GNC Women’s Ultra Mega Energy & ...

My second blue dot 😊 #bluedotchallenge
5 likes
11 hours ago

**Weekly Remaining**
41
Daily Remaining
9
Daily Used

**Breakfast total 3**
Trader Joe's Refrigerated Organic soy milk - vanilla, Coffee, black, without sugar

**Lunch total 11**
Cheese, gorgonzola, Oil, olive, Garlic, fresh, Fage Total 0% Plain nonfat Greek strained yogurt, Ar...

Four blue dots this week and working on my fifth! I’m making it a point not to suffer. Hungry last, watermelon! Any zero point food I want and kept the blue dot:). Tracking everything even on the bad days to create this habit. Going to my meeting every week even when “I don’t feel like it”

#bluedotchallenge
#trackingeverything
#meetingmatter
WellnessWins™

A first-of-its-kind program that rewards members for building healthy habits.

You earn “Wins” for:

- Tracking meals (breakfast, lunch, dinner)
- Tracking activity
- Tracking weight
- Attending workshops
Incentives / Gamify

WellnessWins celebrates outcomes

Milestones are rewarded for:

- Weight loss: 5, 10, 25, 50, 75, 100, 125, 150, 175 & 200 lbs
- Goal weight

“I’ve got my keyring with all my little bangles hanging off of it. I love that thing. It might seem stupid but it was just fun to get those rewards along the way, a physical manifestation of your success”
Incentives / Gamify
Incentives / Gamify
WellnessWins

is motivating

We are holding members accountable while making it fun to earn
Incentives / Gamify

First Steps
105w ago

You tracked your first activity. Take note of how that one simple action made you feel.

Share to Connect

10 kilos de perdu
Il y a 8sem.

10 kilos! Vous suivez le programme à la perfection! Vous voyez la différence dans le miroir? Bien joué!

Partager sur Connect

You did it

Congratulations! Reaching your goal proves that you have the determination to accomplish something amazing. WW has your back as you continue to focus on your wellness and maintain your healthy habits.

Share to Connect

Recognition helps motivate
Incentives / Gamify

Recognition helps motivate
Incentives / Gamify

You’ve tracked 3 dinners in a row

30 Bonus Wins earned

What are Wins?
Incentives / Gamify

Your WW membership goes way beyond weight loss
That's why we're now giving you a monthly update on your journey and highlighting members-only promotions and events. Thank you for being a WW member!

Your month in review: August

You've earned 2100 WellnessWins so far - keep it up.
See all rewards →

- Most tracked meal: Breakfast
- Weigh-ins: 3
- Days of tracked activity: 1
Motivation

- Member is doing the work.
- Important for them to remind *themselves* why they are doing this
“Weight Watchers, for example, coaches dieters to use an array of self-deployed situational and cognitive strategies and, in addition, sponsors in-person meetings, communicates social norms, and provides a phone app to track eating and exercise.”

**Table 1. How Strategies for Reducing Self-Control Failure Might Be Applied to Increasing Healthy Eating**

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self-deployed situational strategies</strong></td>
<td></td>
</tr>
<tr>
<td>Commitment devices</td>
<td>Decision maker commits to eat a serving of fruit and vegetables at dinner every night, asks spouse to serve as a referee, and puts money on the line that will be forfeited to spouse in case he or she fails to meet this commitment.</td>
</tr>
<tr>
<td>Temptation bundling</td>
<td>Decision maker listens to a favorite music album only when cooking dinner from scratch (rather than eating fast food).</td>
</tr>
<tr>
<td>Situation modification</td>
<td>Decision maker stocks up on bags of Halloween candy for trick-or-treaters—but only candy that she does not like.</td>
</tr>
<tr>
<td>Behavior therapy</td>
<td>Decision maker works with a therapist, learning to identify triggers that result in junk food binges (e.g., deadlines at work) and also alternatives (e.g., taking a walk) that can meet the same needs (e.g., stress relief).</td>
</tr>
<tr>
<td><strong>Self-deployed cognitive strategies</strong></td>
<td></td>
</tr>
<tr>
<td>Goal setting</td>
<td>Decision maker decides: “I will eat a fruit or vegetable with every meal!”</td>
</tr>
<tr>
<td>Planning</td>
<td>Decision maker plans: “If it is 8 a.m., then I will look in the refrigerator for some fruit to have with my breakfast.”</td>
</tr>
<tr>
<td>Mental contrasting with implementation intentions</td>
<td>Decision maker thinks: “The best outcome of eating healthy is that I will have more energy. The obstacle that stands in the way is that I don’t have time to go shopping. My plan is: ‘If it is Saturday morning, then I will take a nice walk to the grocery store to buy fresh fruit that I’ll then eat.’”</td>
</tr>
<tr>
<td>Psychological distancing</td>
<td>Decision maker reframes situation using third-person perspective: “Angela is hungry and has a choice between a bag of potato chips and an apple. Which should she choose?”</td>
</tr>
<tr>
<td>Mindfulness</td>
<td>Decision maker introspects: “I notice that I’m craving potato chips. I accept that I have this urge. I may or may not act on it.”</td>
</tr>
<tr>
<td>Cognitive therapy</td>
<td>Decision maker works with a therapist, learning to ask, “What thoughts lead me to snack on potato chips in the afternoon? Do I think, ‘I can’t resist junk food. I have no self-control at all! And is that a reasonable thought? Or am I exaggerating?”</td>
</tr>
</tbody>
</table>
Meta-analysis of 90 articles + 96 field experiments (299 effect sizes), average effect of healthy eating nudges of Cohen’s $d=0.23$.

$= 124$ kcal change in a daily intake
Or -7.2%

8 tablespoons sugar / day
Summary

Highly primed experience

<table>
<thead>
<tr>
<th>easy</th>
<th>SmartPoints, FitPoints</th>
</tr>
</thead>
<tbody>
<tr>
<td>available</td>
<td>In your pocket, AMZN, neighborhood</td>
</tr>
<tr>
<td>supportive</td>
<td>community</td>
</tr>
</tbody>
</table>

Highly saliency

<table>
<thead>
<tr>
<th>progress</th>
</tr>
</thead>
<tbody>
<tr>
<td>food, fitness</td>
</tr>
<tr>
<td>tips</td>
</tr>
</tbody>
</table>

Incentives / Recognition at multiple temporal scales

<table>
<thead>
<tr>
<th>Initial</th>
<th>First track, first barcode scan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily</td>
<td>Blue dot</td>
</tr>
<tr>
<td>Weekly</td>
<td>Weight check in</td>
</tr>
<tr>
<td>Continuous</td>
<td>streaks</td>
</tr>
<tr>
<td>Monthly</td>
<td>review</td>
</tr>
<tr>
<td>Event</td>
<td>Wins, milestone</td>
</tr>
<tr>
<td>Annual</td>
<td>Annual review??</td>
</tr>
</tbody>
</table>

Multiple types of recognition

<table>
<thead>
<tr>
<th>Non tangible</th>
<th>Badges, kudos</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peers</td>
<td>Connect</td>
</tr>
<tr>
<td>Goods in kind</td>
<td>wins</td>
</tr>
<tr>
<td></td>
<td>Initial</td>
</tr>
<tr>
<td>---------------</td>
<td>---------</td>
</tr>
<tr>
<td><strong>Food</strong></td>
<td></td>
</tr>
<tr>
<td>Saliency</td>
<td></td>
</tr>
<tr>
<td>Reward &amp; Recognition (R&amp;R)</td>
<td>badges</td>
</tr>
<tr>
<td>Defaults</td>
<td></td>
</tr>
<tr>
<td><strong>Fitness</strong></td>
<td></td>
</tr>
<tr>
<td>Saliency</td>
<td></td>
</tr>
<tr>
<td>R &amp; R</td>
<td>badges</td>
</tr>
<tr>
<td><strong>Weight</strong></td>
<td></td>
</tr>
<tr>
<td>Saliency</td>
<td></td>
</tr>
<tr>
<td>R &amp; R</td>
<td>badges</td>
</tr>
<tr>
<td>Defaults</td>
<td></td>
</tr>
</tbody>
</table>
The Hook

- Trigger
  - External
  - Internal

- Action

- Investment

- Reward
  - Variable
Data products at WW
Data products at WW

**Growth**
- Churn model
- Return model
- LTV models
- Single Member View

**Program**
- Recipe recommender
- Similar recipes
- Auto-tags
- Clustering member foods
- Composite foods

**Social Network: Connect**
- Personalized feed
- Groups search
- Who to follow

**Infrastructure**
- APIs
  - Primrose
A SURVEY OF FOOD RECOMMENDERS

A PREPRINT

Carl Anderson
Weight Watchers International
New York, USA
carl.anderson@weightwatchers.com

September 18, 2018

ABSTRACT

Everyone eats. However, people don’t always know what to eat. They need a little help and inspiration. Consequently, a number of apps, services, and programs have developed recommenders around food. These cover food, meal, recipe, and restaurant recommendations, which are the most common use cases, but also other areas such as substitute ingredients, menus, and diets. The latter is especially important in the area of health and wellness where users have more specific dietary needs and goals.

In this survey, we review the food recommender literature. We cover the types of systems in terms of their goals and what they are recommending, the datasets and signals that they use to train models, the technical approaches and model types used, as well as some of the system constraints.

Keywords: Personalization · Food recommendation · Recommendation systems · Collaborative filtering · Content-based recommenders · Expert systems

Table 1: High-level summary table that highlights the breadth of food recommender space, covering what is being recommended to whom, how, and why. * represents more speculative examples.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Examples</th>
</tr>
</thead>
</table>
| **Who** are the users? | Hungry people: you might like to order this meal  
Cooking enthusiasts: you might like to make this recipe  
Health-conscious: you’ll love this healthy, nutritious lunch  
Dieters: this is a low-calorie but filling and healthy meal  
Patients: doctors suggest that you follow this diet |
| **What** is being recommended? | Ingredient: you can substitute butter with sour cream for reduced fat and calories  
Food: we think you’ll like these summer rolls  
Meal: we think you’ll like this chicken breast plate with rice and broccoli  
Recipe: try this pecan pie recipe  
Recipe collection: here is a set of salad recipes you’ll love  
Restaurant: you have to try Danny’s Pizza  
Cuisine*: as you like Thai, you might like Indonesian food too  
Diet / menu / meal plan: this is a low-sodium diet that ought to work for you |
| **When** is it being recommended? | Realtime: where should I eat now; what’s near me?  
Batch: here is your weekly email of recipes, just for you |
| **Why** is it being recommended? | Taste: here is something you might like to eat / make / order  
Health & wellness: to help people become or remain healthy, to help people lose weight, and to help patients recover |
Food is at the core of our product.
Recipe Recommendations

Similar Recipes

Dinner Recommendations

Because you tracked **Chicken-Fried Steak**

- Chicken Fried Rice
- Southern-Style Oven-Fried Chicken
- Vegetable Fried Rice
- Stir-Fried Chili Mango Chicken with Peppers

EASY	MODERATE	EASY	MODERATE	EASY

Because you tracked **Grilled Salmon with Mustard-Herb Crust**

- Grilled Salmon with Caper Chimichurri
- Honey-Mustard Roasted Salmon
- Lemon-Herb Roasted Salmon
- Salmon Cakes with Dijon-Herb Sauce
- Grilled Tuna with Mustard-Dill Sauce

EASY	EASY	EASY	EASY	MODERATE

# note: push tokenization and and handling of ngrams down to tokenize in concrete classes

```
self.tfidf = TfidfVectorizer(tokenizer=self.tokenize)
self.term_document_matrix = self.tfidf.fit_transform(self.docs)
def cosine_similarity_matrix(self):
    return cosine_similarity(self.term_document_matrix)
```
Similar Recipes Flow

US WW Recipes

*Only recipes with images*

Similar Ingredients

document = ingredient list or name string

lemmatize, tokenize, TF-IDF

Cosine similarity

Rank

Similar Names

Filters

dietary
course
cuisine
main ingredient

Wellness that Works.
Dinner Recommendations Flow

US WW Recipes

Similar Ingredients

Similar Names

Business Logic

Eligible Members

2 weeks of tracking history
Tracked >= 1 recipe
US members

Potential Recs

tracked
most similar
2nd most sim.

n = 4 recommendations

Wellness that Works.
Food Embeddings

- **Motivation**: want to learn a space of foods where similar foods are located near each other
- **Applications**
  - Recommend low point substitute foods
  - Input into recipe recommender
  - Classify new foods and users
- **How to do this?** Word embeddings!
Word Embedding Overview

- Dense real-valued vectors representing word meaning
- **Idea:** words with similar meanings are grouped together in the embedding space
- Many forms of meaning are conflated since there is only one representation per word
FastText Behind the Scenes

- Learns embeddings using either Skip-gram or CBOW algorithms
- But learns representations for sub-word units rather than entire words
- Representations for whole words are composed from subword representations

Wellness that Works.
Preliminary Attempts

- **Attempt 1 (using food log data)** [1]:
  - Split food names into tokens
  - Each food name = 1 document
  - Average token embeddings for food name
  - Append calorie-normalized nutritional info
  - Did not work well, but might work with better preprocessing

- **Attempt 2 (using recipe data)** [2]:
  - Context = recipe ingredients
  - Each recipe = 1 document
  - Did not work well, recipe data too small

[2] Cooking up Food Embeddings; Sauer et. al 2017
Final Attempt

- Context = ordered food entries, grouped by user id and time of day (meal type) over one week
- Preprocessed data same way as in [2]
- Each “word” in a document is a whole food name
- Best result from using subword unit modeling
- Other ideas: filtering for power users

(UID=1234, breakfast, week1) = [Monday breakfast, Tuesday breakfast, Wednesday breakfast,...]

= [coffee, toast, jam, apple, coffee, orange_juice, tea, cereal, 2%_milk, banana,...]

Will learn associations among items:

within meal: cereal↔2% milk, cereal↔whole milk
among meals: apple↔banana, coffee↔tea
One of the main goals of the project was to extract substitute food items.

Food data contains category information.

Simply eliminate results from NN list that are not in the same category.
Personalizing Social Network
#wednesdaymotivation

Down 54.6 lbs!
Never going back to that daddy again!
If you need to re... See More

2 days ago

537 likes 64 comments

rockerroad99

#throwbackthursday Italy edition 🇮🇹

I was 25 years younger and over 100 pounds heavier back in 1994 when I visited Italy. I remember how hard it was to get... See More

1 day ago

926 likes 70 comments

auntmarci

#facetofacefriday Italy edition

And look I have a neck, too!!

25 years ago i was 100 pounds hea... See More

8 hours ago

373 likes 22 comments

sabinec58
Seeking positivity
Getting help
Sharing goals
Encouraging others
Making friends
Building a brand

Qualitative research from our experience research / analytics teams
I want to feel good and see inspirational and useful posts.
I post when I have questions or need encouragement.
I post to show what I did or am going to do.
Encouraging others

I give back the support I received to those who need it.
I build and invest in meaningful relationships.
I create content to grow a large following.
“Hidden” agendas may change throughout the day or over the course of a member’s journey
### Personalized Feed

#### Video Feed Options:
- Video
- Video
- Content-based
- Before / After

#### Feed Content Types:
- Collaborative filter
- Content-based
- Videos
- Before / After
- Popular

**Wellness that Works:**
Personalized Feed

The Multi-Armed Bandit Problem

- We have $d$ arms. For example, arms are ads that we display to users each time they connect to a web page.
- Each time a user connects to this web page, that makes a round.
- At each round $n$, we choose one ad to display to the user.
- At each round $n$, ad $i$ gives reward $r_i(n) \in \{0, 1\}$: $r_i(n) = 1$ if the user clicked on the ad $i$, $0$ if the user didn’t.
- Our goal is to maximize the total reward we get over many rounds.

Thompson Sampling Algorithm

Step 1. At each round $n$, we consider two numbers for each ad $i$:
- $N_i^1(n)$ - the number of times the ad $i$ got reward $1$ up to round $n$,
- $N_i^0(n)$ - the number of times the ad $i$ got reward $0$ up to round $n$.

Step 2. For each ad $i$, we take a random draw from the distribution below:

$$\theta_i(n) = \beta(N_i^1(n) + 1, N_i^0(n) + 1)$$

Step 3. We select the ad that has the highest $\theta_i(n)$. 

Wellness that Works.
**Example:** I love biking. Is there a group for this?

**Issue:** search only uses title and description, not content

**Solution:** we’ve provided top 100 terms + top 100 hashtags per group

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Camping</th>
<th>Brides</th>
</tr>
</thead>
<tbody>
<tr>
<td>rider</td>
<td>campground</td>
<td>bride</td>
</tr>
<tr>
<td>cyclist</td>
<td>tent</td>
<td>bridal</td>
</tr>
<tr>
<td>cycling</td>
<td>camping</td>
<td>alteration</td>
</tr>
<tr>
<td>ally</td>
<td>camper</td>
<td>bridesmaid</td>
</tr>
<tr>
<td>century</td>
<td>trailer</td>
<td>venue</td>
</tr>
<tr>
<td>biking</td>
<td>teardrop</td>
<td>seamstress</td>
</tr>
<tr>
<td>output</td>
<td>campfire</td>
<td>engagement</td>
</tr>
<tr>
<td>cadence</td>
<td>camped</td>
<td>ceremony</td>
</tr>
</tbody>
</table>
Who To Follow
Computer Science > Information Retrieval

Metadata Embeddings for User and Item Cold-start Recommendations

Maciej Kula

(Submitted on 30 Jul 2015)

I present a hybrid matrix factorisation model representing users and items as linear combinations of their content features' latent factors. The model outperforms both collaborative and content-based models in cold-start or sparse interaction data scenarios (using both user and item metadata), and performs at least as well as a pure collaborative matrix factorisation model where interaction data is abundant. Additionally, feature embeddings produced by the model encode semantic information in a way reminiscent of word embedding approaches, making them useful for a range of related tasks such as tag recommendations.

Subjects: Information Retrieval (cs.IR)
ACM Classes: H.3.3
Cite as: arXiv:1507.08439 [cs.IR]
(or arXiv:1507.08439v1 [cs.IR] for this version)
Who To Follow

Features:
- Demographics: age, gender...
- Location
- Membership: type...
- Goal / Weight
- Tags, interactions
- Groups

Wellness that Works.
Who To Follow

- Trigger Cron Job
  - Nightly Batch Training
  - Data Lake

- Serialized models
  - Cloud Storage

- Cached features
  - Redis Cache

- Trigger Deployment
  - Live Prediction API

- On train success, kick off redeployment of API service to pick up new objects

Client

Wellness that Works.
Primrose
Taking Stock of our own challenges

What would make a good recommender system at WW?

Slow serialization

but our medium data can be kept in RAM...

No live features

but we know Docker, k8s...

Easy onboarding

mono repo with config as code...
Primrose has features to address each design consideration

**Primrose:** *(Production In-Memory Solution)* framework for solving WW’s most common use cases, caching batched predictions with machine-learning engineering baked-in.

- **Data science**
  - Python *in-memory DAG* runner, with no serialization between nodes of the DAG.
- **Infrastructure**
  - DAG is defined as *configuration-as-code* approach -- one container for all models
- **People**
  - Abstract ML and data manipulation operations, data scientists can easily extend the framework
Primrose: a framework for simple, quick modeling deployments
and we open sourced it....
**Primrose 1.0.6**

**pip install primrose**

Primrose: a framework for simple, quick modeling deployments

**Navigation**
- Project description
- Release History
- Download files

**Project links**
- Homepage
- Source
- Documentation

**Statistics**
GitHub statistics:
- Stars: 9
- Forks: 0
- Open issues/Pulls: 1

View statistics for this project via Libraries.io or by using Google Bigquery.

**Meta**

---

**Overview**

**Project description**

**Primrose at a glance**

Primrose is a simple Python framework for executing in-memory workflows defined by directed acyclic graphs (DAGs) via configuration files. Data in primrose flows from one node to another while avoiding serialization, except for when explicitly specified by the user. Primrose nodes are designed for simple batch-based machine learning workflows, which have datasets small enough to fit into a single machine's memory.

**Table of Contents**

We suggest reading the documentation in the following order:

- **Overview and motivation** for primrose—this file.
- **Getting Started**—run your first primrose jobs.
- **DAG Configurations**—primrose adopts a configuration-as-code paradigm. This section introduces primrose configuration files.
- **Metadata**—this covers more advanced options of the configuration files.
- **Command Line Interface (CLI)**—run commands using the CLI.
- **Developer Notes**—how to create your own node classes.
- **DataObjects**—a deep dive into DataObjects, the core data handling and bookkeeping object.
Primrose jobs are executed as Directed Acyclic Graphs (DAG)s in python

**Flexibility:** any number of operations allowed in a single DAG, across any python library

**Data and functions are passed between nodes** in an object that understands how to extract the correct data for each node.
DAGs are composed of implementation agnostic, extensible nodes for data science

Data scientists can write individual nodes using any Python framework or library they choose.
Primrose is run like an ETL pipeline in a single docker container for each configuration.
For simpler deployments: Primrose uses a “configuration as code” approach

Object configuration and DAG structure are build in a configuration JSON

Recipe recommender DAG JSON

Churn Model DAG JSON

Connect Feed DAG JSON

Primrose validates the configuration and instantiates the correct classes at runtime

Different outputs and results for each DAG

Success, fame, money...
Primrose config snippet: cluster with KMeans

```
"kmeans_cluster_model":{
  "class": "SklearnClusterModel",
  "mode": "train",
  "features": ["x1","x2"],
  "model": {
    "class": "cluster.KMeans",
    "args": {"n_clusters": 6, "random_state": 42}
  },
  "destinations": ["write_data", "write_model"]
}
```

Primrose config snippet: use DBSCAN instead

```
"dbscan_cluster_model":{
  "class": "SklearnClusterModel",
  "mode": "train",
  "features": ["x1","x2"],
  "model": {
    "class": "cluster.DBSCAN",
    "args": {"min_samples": 3}
  },
  "destinations": ["write_data", "write_model"]
}
```
Similar Recipes Flow

US WW Recipes
*Only recipes with images*

Similar Ingredients

document = ingredient list or name string

lemmatize, tokenize, TF-IDF

Cosine similarity

Rank

Filters
dietary
course
cuisine
main ingredient
Productionalize in Primrose DAG

- Google BigQuery Data lake Reader
- NLTK + Custom Lemmatization
- Sklearn TF-IDF + cosine similarity
- Write to GCS Bucket and Google MemoryStore

Success!

logging.info('Your newbie DS has written production quality code.'
Productionalize in Primrose DAG

Google BigQuery Data lake Reader

NLTK + Custom Lemmatization

Sklearn TF-IDF + cosine similarity

```
"ingredient_model": {
  "class": "RecipeIngredientSearchEngine",
  "mode": "predict",
  "id_key": "recipeID",
  "doc_key": "ingredient_string",
  "destinations": [
    "recipe_postprocess"
  ]
},

"name_model": {
  "class": "RecipeNameSearchEngine",
  "mode": "predict",
  "id_key": "recipeID",
  "doc_key": "displayName",
  "destinations": [
    "recipe_postprocess"
  ]
}
```

```
"corpus_pipeline": {
  "class": "SimilarRecipesPipeline",
  "nonfood_image_exclude_list": "data/non_food_image_exclude_list.txt",
  "filter_out_non_image": true,
  "remove_dupe_ingredients": true,
  "filter_in_dinner": false,
  "is_training": true,
  "popularity_scaling_min": 0.4,
  "popularity_scaling_max": 0.6,
  "destinations": [
    "ingredient_model",
    "name_model"
  ]
}
```
Productionalize in Primrose DAG

Google BigQuery Data lake Reader

NLTK + Custom Lemmatization

Sklearn TF-IDF + cosine similarity

Business Logic (filters)

Write to GCS Bucket and Google MemoryStore

Success!
logging.info('Your newbie DS has written production quality code.')
Dinner Recommendations Flow

**US WW Recipes**

**Similar Ingredients**

**Similar Names**

**Eligible Members**

2 weeks of tracking history
Tracked >= 1 recipe
US members

**Potential Recs**

tracked
most similar
2nd most sim.

**Business Logic**

n = 4 recommendations

Wellness that Works.
Productionalizing is easier the second time

Same BQ reader class, different SQL input file

New postprocess class to sort, filter and interleave potential recommendations

Success!

logging.warning('Data Scientist is developing software engineering skills.')
Primrose has features to address each design consideration

**Primrose**: *(Production In-Memory Solution)* framework for solving WW’s most common use cases, caching batched predictions with machine-learning engineering baked-in.

- **Data science**: Python in-memory DAG runner, with no serialization between nodes of the DAG.
- **Infrastructure**: DAG is defined as configuration-as-code approach -- one container for all models
- **People**: Abstract ML and data manipulation operations, data scientists can easily extend the framework
Wrap Up

Nudges:
- Once is not enough: nudge different times, channels, timescales
- Recognition really important: Nudge before, recognize after
- Holistic view: challenges, community, personality

Primrose:
- In-memory, config-as-code, extensible
- Helped our new team be productive and get models into prod
- Available today
Questions

- carl.anderson@ww.com
- @leapingllamas

- Food RecSys: https://arxiv.org/abs/1809.02862
- Primrose: https://github.com/ww-tech/primrose
- Tech blog: https://medium.com/ww-tech-blog

Hiring: especially data scientists in Toronto