Big Data & AI in Business

How to become a data driven business

Session 5: Value Identification in different industries

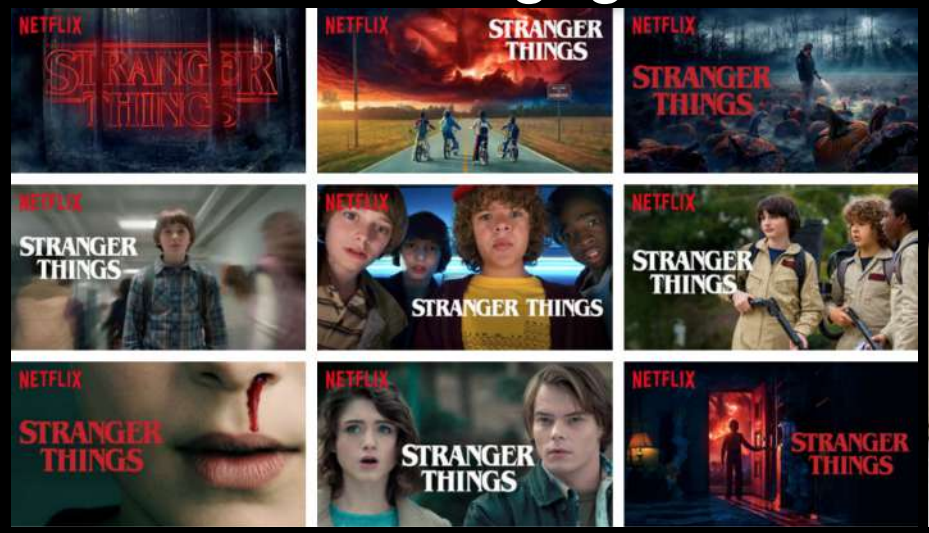
David **G Pisano**

How to identify Value

What is the value of AI for business?

New ways of working

Customer Engagement



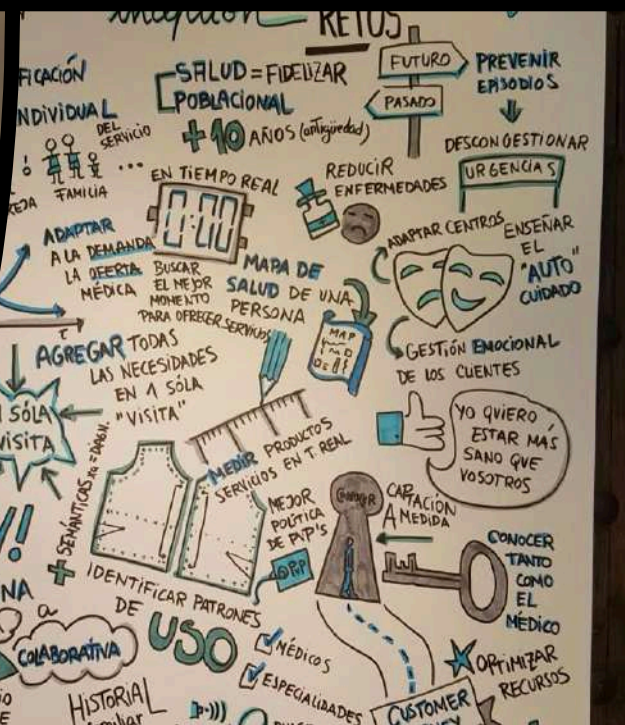
Operational Efficiency

New Businesses

Discovery Workshops

1	2	3	4	5
Preparation & Introduction	Problem selection	Use case definition	Priorization Matrix	C-level presentation

Human center design	AI center designed
Business oriented	Data oriented
Fall in love with the problem	Fall in love with the data
People from all levels	Top managers
Give context: what is AI?	Everybody knows what is AI
Pick your product owner and your sponsors. This is your new community	Thank you for your ideas now are mine
Prioritize and select top 5	We will make an analysis…
Visual is everything…	This about maths…



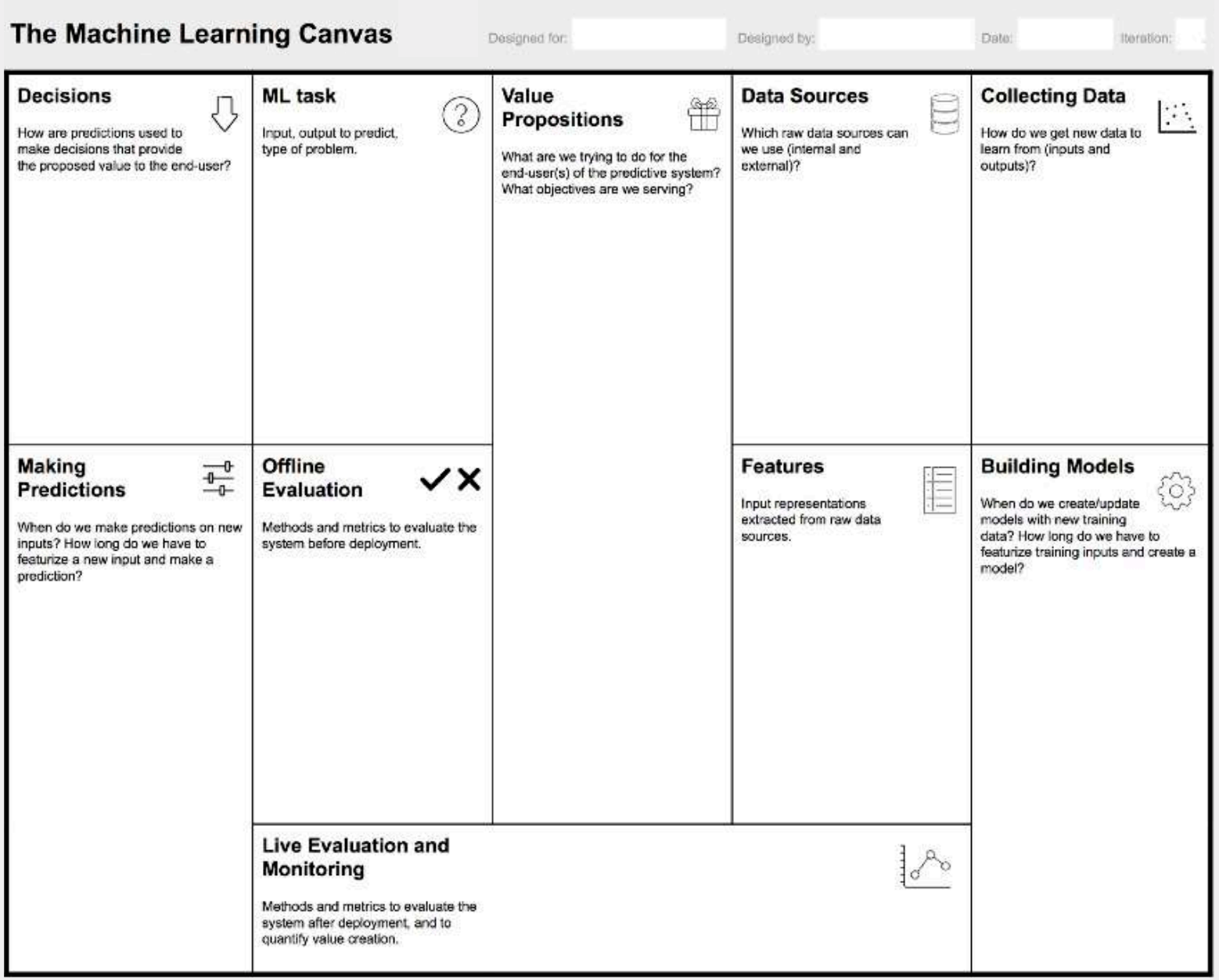
How to define AI use cases

THE

MACHINE LEARNING CANVAS

A handbook for innovators and visionary managers striving to design tomorrow's Machine Learning systems

LOUIS DORARD, PH.D.



The Machine Learning Canvas

Designed for: | Designed by: | Date: | Iteration:

Decisions

How are predictions used to make decisions that provide the proposed value to the end-user?

ML task

Input, output to predict, type of problem.

Value Propositions

What are we trying to do for the end-user(s) of the predictive system? What objectives are we serving?

Data Sources

Which raw data sources can we use (internal and external)?

Collecting Data

How do we get new data to learn from (inputs and outputs)?

Making Predictions

When do we make predictions on new inputs? How long do we have to featurize a new input and make a prediction?

Offline Evaluation

Methods and metrics to evaluate the system before deployment.

Features

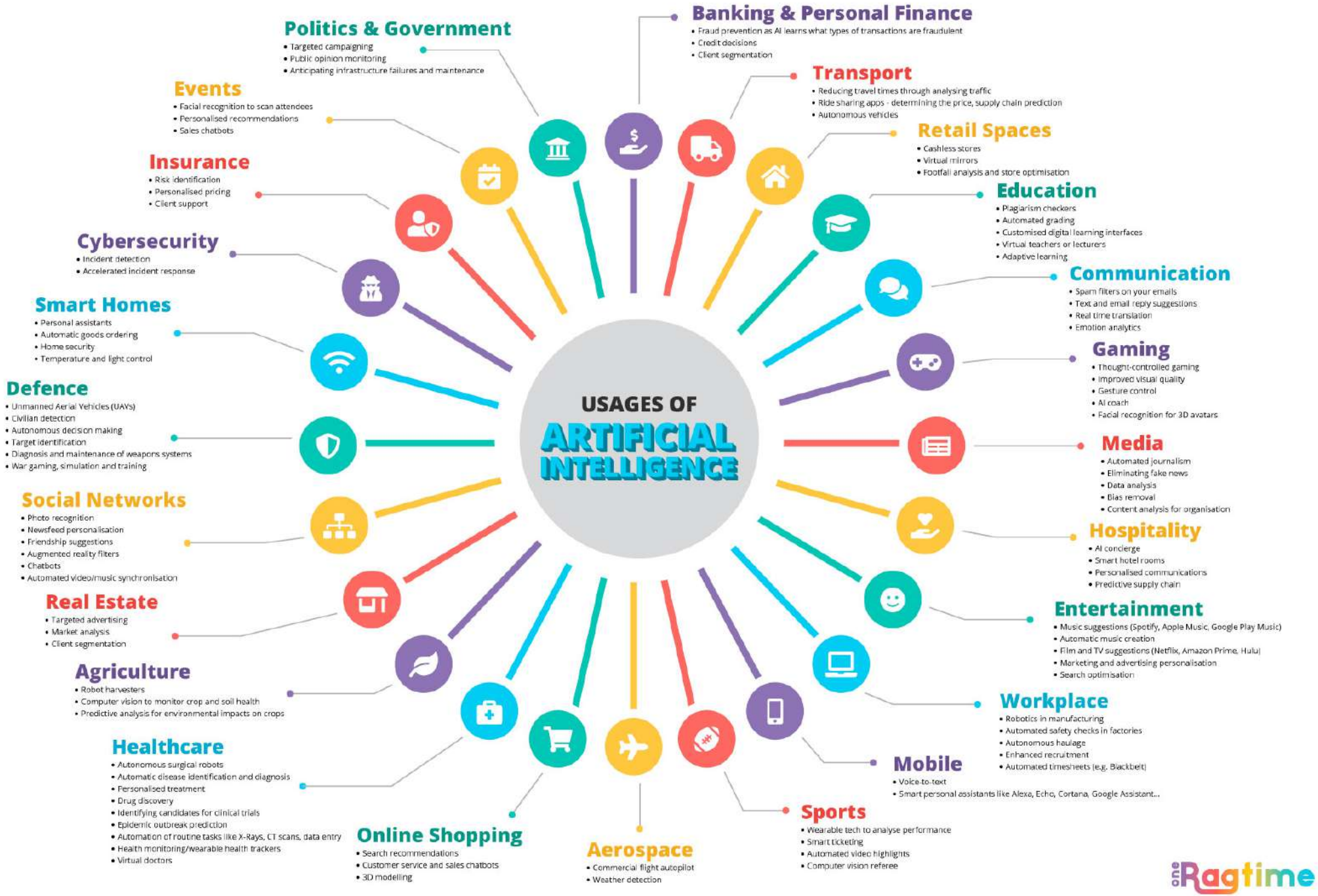
Input representations extracted from raw data sources.

Building Models

When do we create/update models with new training data? How long do we have to featurize training inputs and create a model?

Live Evaluation and Monitoring

Methods and metrics to evaluate the system after deployment, and to quantify value creation.


USAGES OF
ARTIFICIAL INTELLIGENCE
Politics & Government
• Targeted campaigning
• Public opinion monitoring
• Anticipating infrastructure failures and maintenance
Banking & Personal Finance
• Fraud prevention as AI learns what types of transactions are fraudulent
• Credit decisions
• Client segmentation
Events
• Facial recognition to scan attendees
• Personalised recommendations
• Sales chatbots
Transport
• Reducing travel times through analysing traffic
• Ride sharing apps - determining the price, supply chain prediction
• Autonomous vehicles
Insurance
• Risk identification
• Personalised pricing
• Client support
Retail Spaces
• Cashless stores
• Virtual mirrors
• Footfall analysis and store optimisation
Cybersecurity
• Incident detection
• Accelerated incident response
Education
• Plagiarism checkers
• Automated grading
• Customised digital learning interfaces
• Virtual teachers or lecturers
• Adaptive learning
Smart Homes
• Personal assistants
• Automatic goods ordering
• Home security
• Temperature and light control
Communication
• Spam filters on your emails
• Text and email reply suggestions
• Real time translation
• Emotion analytics
Defence
• Unmanned Aerial Vehicles (UAVs)
• Civilian detection
• Autonomous decision making
• Target identification
• Diagnosis and maintenance of weapons systems
• War gaming, simulation and training
Gaming
• Thought-controlled gaming
• Improved visual quality
• Gesture control
• AI coach
• Facial recognition for 3D avatars
Social Networks
• Photo recognition
• Newsfeed personalisation
• Friendship suggestions
• Augmented reality filters
• Chatbots
• Automated video/music synchronisation
Media
• Automated journalism
• Eliminating fake news
• Data analysis
• Bias removal
• Content analysis for organisation
Real Estate
• Targeted advertising
• Market analysis
• Client segmentation
Hospitality
• AI concierge
• Smart hotel rooms
• Personalised communications
• Predictive supply chain
Agriculture
• Robot harvesters
• Computer vision to monitor crop and soil health
• Predictive analysis for environmental impacts on crops
Entertainment
• Music suggestions (Spotify, Apple Music, Google Play Music)
• Automatic music creation
• Film and TV suggestions (Netflix, Amazon Prime, Hulu)
• Marketing and advertising personalisation
• Search optimisation
Workplace
• Robotics in manufacturing
• Automated safety checks in factories
• Autonomous haulage
• Enhanced recruitment
• Automated timesheets (e.g. Blackbelt)
Healthcare
• Autonomous surgical robots
• Automatic disease identification and diagnosis
• Personalised treatment
• Drug discovery
• Identifying candidates for clinical trials
• Epidemic outbreak prediction
• Automation of routine tasks like X-Rays, CT scans, data entry
• Health monitoring/wearable health trackers
• Virtual doctors
Mobile
• Voice-to-text
• Smart personal assistants like Alexa, Echo, Cortana, Google Assistant...
Online Shopping
• Search recommendations
• Customer service and sales chatbots
• 3D modelling
Aerospace
• Commercial flight autopilot
• Weather detection
Sports
• Wearable tech to analyse performance
• Smart ticketing
• Automated video highlights
• Computer vision referee

1 AI Value Identification

Understanding AI benefits & use cases is the main challenge to AI/ ML Adoption for 42% of executives

Gartner

1

2

3

Understand what AI is: Identifying various disciplines and techniques is critical to set the right expectations with business

Ideation phase to collect a set of use cases, which may arise from existing business processes or by looking at competitors or other industries

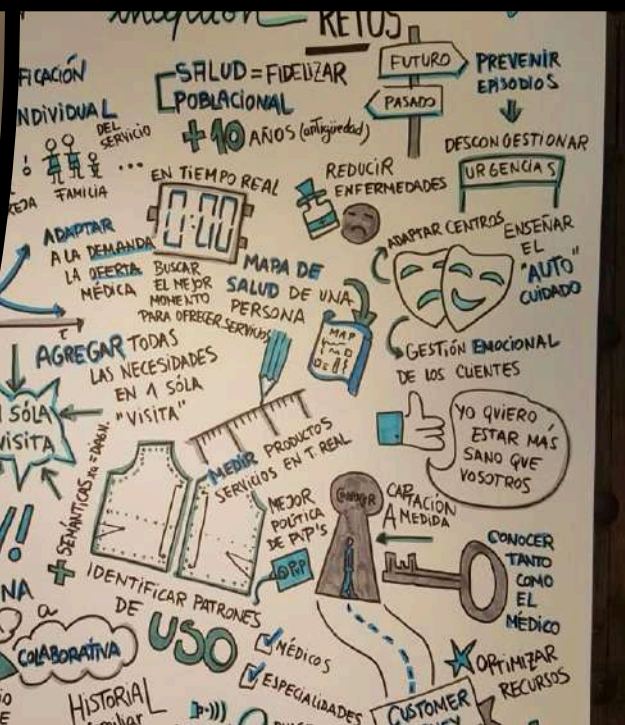
Set your ambition:

- How much do we think AI can transform business?
- How strategic is AI for business?
- Who is leading the value identification process?

Discovery Workshops

1	2	3	4	5
Preparation & Introduction	Problem selection	Use case definition	Priorization Matrix	C-level presentation

Human center design	AI center designed
Business oriented	Data oriented
Fall in love with the problem	Fall in love with the data
People from all levels	Top managers
Give context: what is AI?	Everybody knows what is AI
Pick your product owner and your sponsors. This is your new community	Thank you for your ideas now are mine
Prioritize and select top 5	We will make an analysis…
Visual is everything…	This about maths…



THE

MACHINE LEARNING CANVAS

A handbook for innovators and visionary managers striving to design tomorrow's Machine Learning systems

LOUIS DORARD, PH.D.

The Machine Learning Canvas (v0.4)

Designed for: | Designed by: | Date: | Iteration:

Decisions

How are predictions used to make decisions that provide the proposed value to the end-user?

ML task

Input, output to predict, type of problem.

Value Propositions

What are we trying to do for the end-user(s) of the predictive system? What objectives are we serving?

Data Sources

Which raw data sources can we use (internal and external)?

Collecting Data

How do we get new data to learn from (inputs and outputs)?

Making Predictions

When do we make predictions on new inputs? How long do we have to featurize a new input and make a prediction?

Offline Evaluation

Methods and metrics to evaluate the system before deployment.

Features

Input representations extracted from raw data sources.

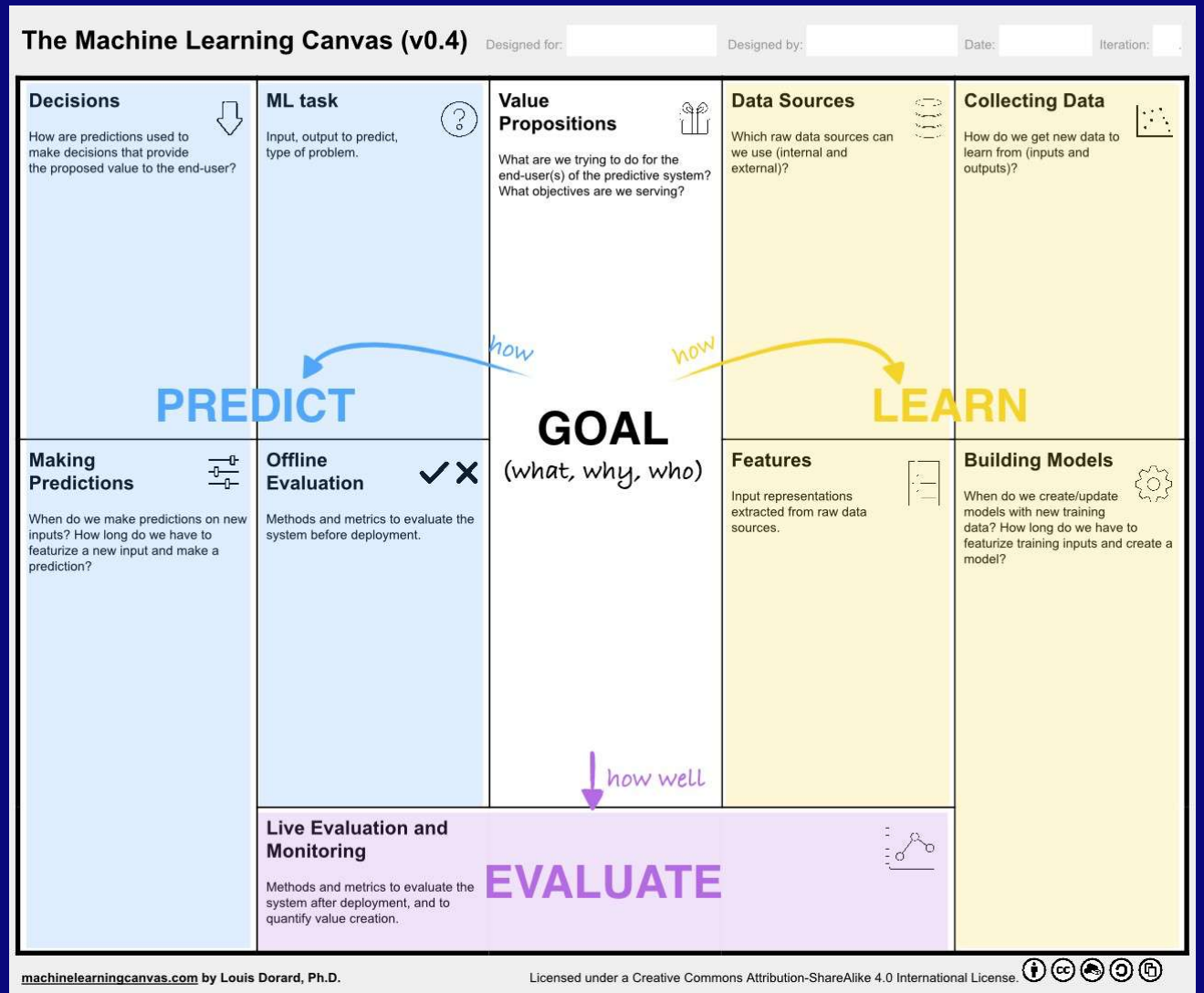
Building Models

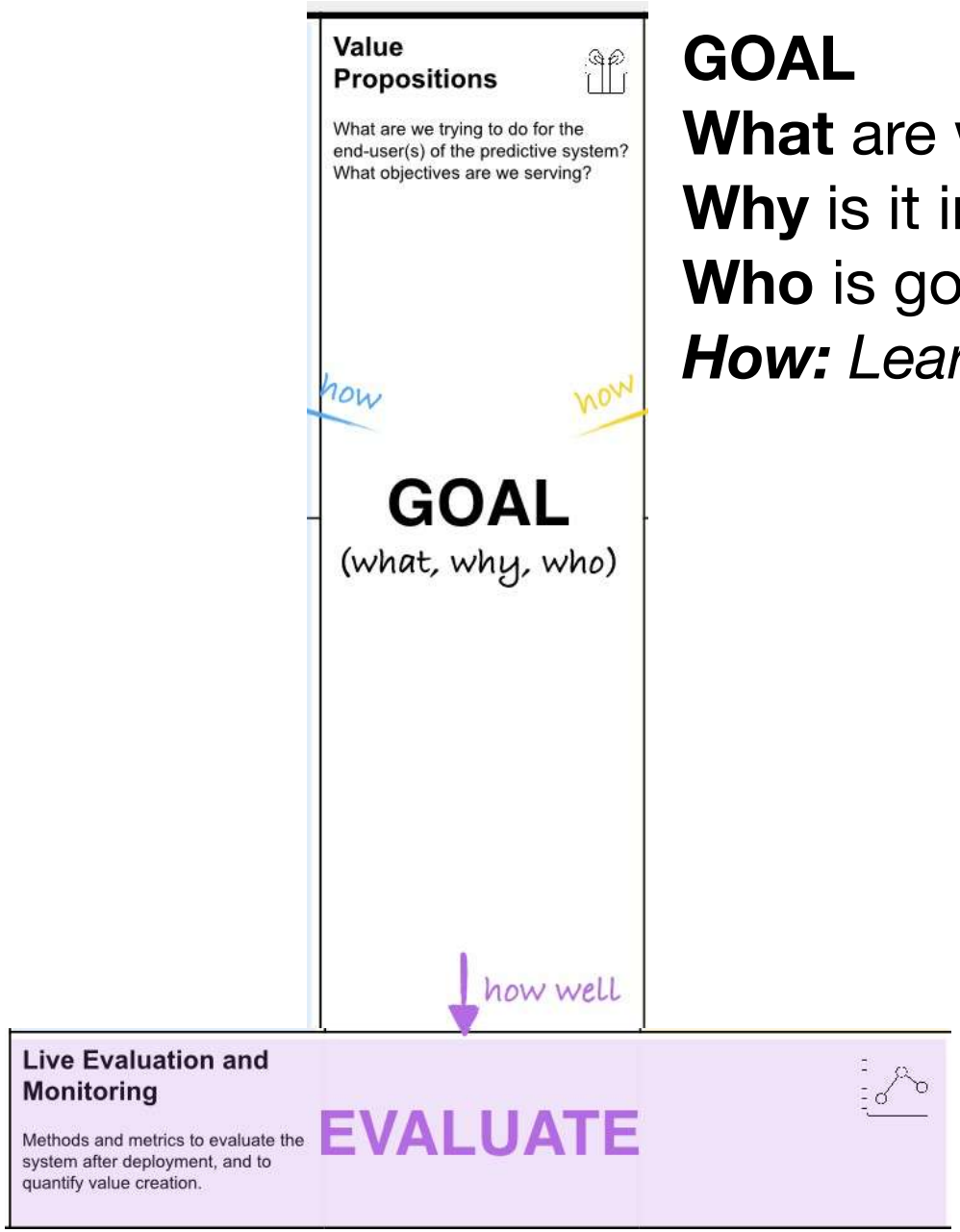
When do we create/update models with new training data? How long do we have to featurize training inputs and create a model?

Live Evaluation and Monitoring

Methods and metrics to evaluate the system after deployment, and to quantify value creation.

- What data are we learning from
- How are we using predictions powered by that learning
- How are we making sure that the whole thing “works” through time?





GOAL

What are why trying to do?

Why is it important for the business?

Who is going to use the system/be impacted by it?

How: *Learning & Predictions*

Volume

How much money?

Time to Value? (Velocity)

How long does it take to start to collect that value?

Veracity

How can the company check we are really capturing all that value?

address	bedrooms	bathrooms	size_sqft	lot_size	price
7103 Wolf Rivers Ave Las Vegas NV 89131	4	4	3,811	13,939	495,000
10669 Oak Crest Ave Las Vegas NV 89144	3	2	1,622	5,662	240,000
128 Celia Pl Las Vegas NV 89145	4	2	1,895	6,534	165,000
517 Carpenter Dr Las Vegas NV 89107	4	2	1,286	6,534	120,000

LEARNING from data

Data sources

"Which raw data sources can we use? (internal and external)"

The actual data to be fed to ML algorithms, which will be extracted from sources listed here. For instance, you could be using different types of **internal/external databases, APIs, files, web scraping**, etc.

Data collection

"How do we get new data to learn from? (inputs AND outputs)"

Getting example outputs can often be a barrier to using ML, so this is what you should think about first! There can be a cost associated to getting output data, you need humans to manually look at example reviews and assign them a fake or real label (typically via a human-powered API like that of MTurk or CrowdFlower, which have straightforward cost structures that allow you to anticipate how much data you can afford)

Feature Engineering

"*How do we transform raw data sources into training input representations?"*

This is the art of Machine learning: Invent features capable to predict. Raw data rarely predicts… you'll use features that are aggregates of the raw data, and you should think about what the aggregation is over. For instance, for spacial data you should determine a radius. For temporal data, you should specify a period of time; for instance, one way to represent a customer could be with the number of times they used the service in the last X months. It's useful to highlight such parameters as they can impact the performance of the whole system

Building/Updating models

When do we create/update models with new training data? How long/often do we have to featurize training inputs and create a model?

There are two main reasons why you would want to update models: because having more data could lead to better models, or because the dynamics of whatever phenomenon or behavior you're capturing with data might have changed.

PREDICT

ML Task

"What is the input, the output to predict and the type of problem?"

It helps to define ML tasks as questions about a certain "object" of the real world (which we call the input). The question to answer must be specific, for instance: "Is this email important?" or "How much is this property worth?" The attributes/ characteristics of inputs that allow us to represent them in a computer program, would be listed in the Features box. The output is the answer to the question: all the possible outputs for a classification task (important/spam), or their range for a regression task (10 USD - 10M USD), and their distribution.

Decisions

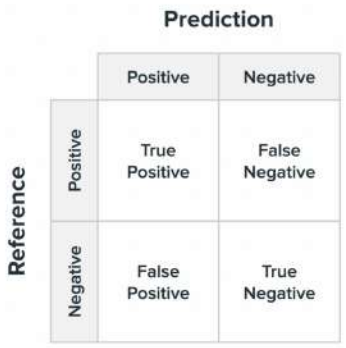
"When and how are predictions used to make decisions that provide the proposed value to the end-user?"

Asking yourself "What if I had perfect predictions?" makes it easier to think about these decisions, before spending too much time on model building.

Making predictions

"When do we make predictions on new inputs? How long do we have to featurize new input and make a prediction?"

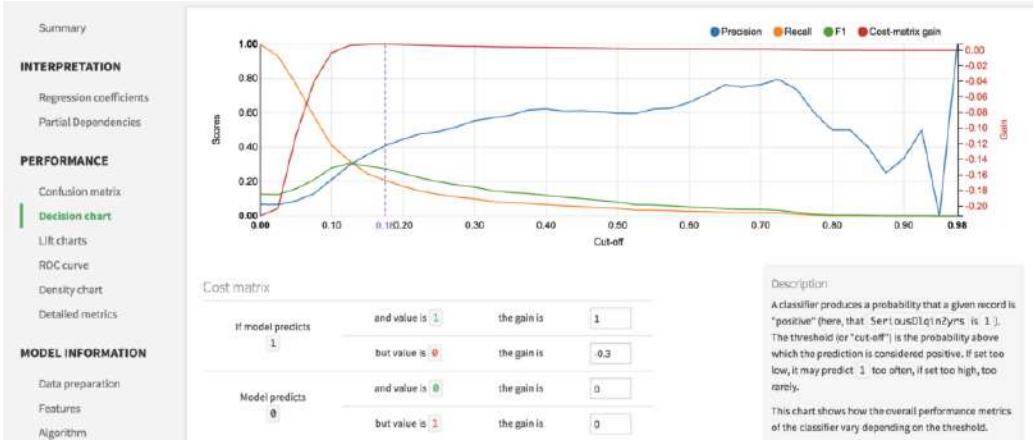
Technical constraints on predictions made to support decisions: volume, frequency, time, etc.



Offline evaluation

"What methods and metrics do we use to evaluate the system before deployment?"

Is this model good enough for my goal? The art of testing



Sizmek's AI-powered platform provides insights that help companies better understand customer data, and consequently, produce more relevant content and campaigns. Sizmek's AI technology analyzes billions of data points (Supervised ML) to help predict the best times, messages and environments for increasing conversion.

Decisions

How are predictions used to make decisions that provide the proposed value to the end-user?

- Filter out ‘no clicks’
- Solicit as many customers as suggested by simulation
- Randomly filter specific labeled data by x % (hold-out rate)

ML task

Input, output to predict, type of problem.

- Predict if a customer will click or not.
- Input: Customer
- Outputs: “click” or “no click” class
- Binary Classification

Value Propositions

What are we trying to do for the end-user(s) of the predictive system? What objectives are we serving?

- Brings together the technology, intelligence, creative solutions, and strategic service for you to create inspiring, seamless advertising that optimizes your marketing budgets and cultivates deeper relationship with your customers around the world.

Data Sources

Which raw data sources can we use (internal and external)?

- Payments database
- Customer support
- CRM Tool
- Emails
- Campaign conversion rate

Collecting Data

How do we get new data to learn from (inputs and outputs)?

- Identify which of previous campaign/ customers churned or not by looking through the payment database.

Making Predictions

When do we make predictions on new inputs? How long do we have to featurize a new input and make a prediction?

- Automatically identify and feature all existing and past customers and make predictions for the current clients.

Offline Evaluation

Methods and metrics to evaluate the system before deployment.

- Evaluate new model’s accuracy on pre-defined customer profiles
- Simulate decision taken on previous months’ customers and its contents.

Features

Input representations extracted from raw data sources.

- Customer support interactions
- Basic customer information (age, city, ethnicity, etc)
- Usage of product: number of logins, functionality used, etc.

Building Models

When do we create/update models with new training data? How long do we have to featurize training inputs and create a model?

- Every marketing campaign, create a new model from the previous campaigns’ hold-out set.

Live Evaluation and Monitoring

Methods and metrics to evaluate the system after deployment, and to quantify value creation.

- Compare churn rate and lost revenue between previous hold-out sets and remaining sets.
- Evaluate the accuracy of previous prediction on hold-out rate
- Monitor ROI (comparing previous campaigns)

Supervised Machine Learning:

Corals, which are colonial living animal, just before dying and leaving the exoskeleton, they emit a fluorescent color, a last chemical "sunscreen" to protect from the excessive heat waves which bring them to bleach.

Scientists have positioned across the world devices that can ping lights and detected remotly the mirrored returning wavelength as spectral signal.

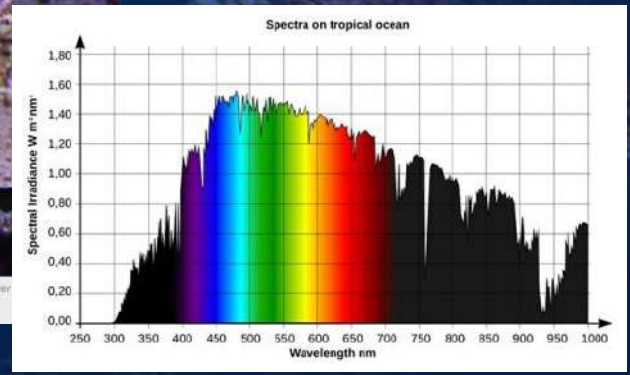
They trained the machine to recognize patterns among these spectral diagrams to understand the reef's stress level in order to predict major future bleaching events.

Ping Coral Reef with light to predict their stress level.

CoralspeQ understand photosynthetic parameters derived from the known absorbance and fluorescence measurements emitted from corals.

David Kramer

The 'CoralSpeq' sensor measures the photosynthetic activity of coral reefs.



https://www.nature.com/news/computers-on-the-reef-1.20497

Supervised Machine Learning: Corals spectrography

Decisions

Compute predictions to help governments to evaluate the establishment of marine protected areas which would help reef's resilience against bleach effects.

ML task

Are these corals bleaching?:
Imput: Spectrogram
Output: Value

-Classification Task given a train set of known spectrograms

Value Proposition

Monitor the health of the reefs worldwide and predict major coral bleach events caused by heatwaves by ping light to coral and mesure the spectrography diagrams.

Data Sources

-Open Source Data
-Google Maps
-US National Oceanic and Atmospheric Administration

Collecting Data

Daily

300 sensors in 18 countries

Making Predictions

The goal is to understand the color patterns of coral before they start to bleach. Predicting which condition bring coral to bleach will help to protect proactivly.

Offline Evaluation

Fish density

Money Spent

Features

-Spectograms

Building Models

Build the training set with physical dives and field data acquisition

Updating the model monthly with the new results to have a richer training set

Live Evaluation and Monitoring

MPAs established

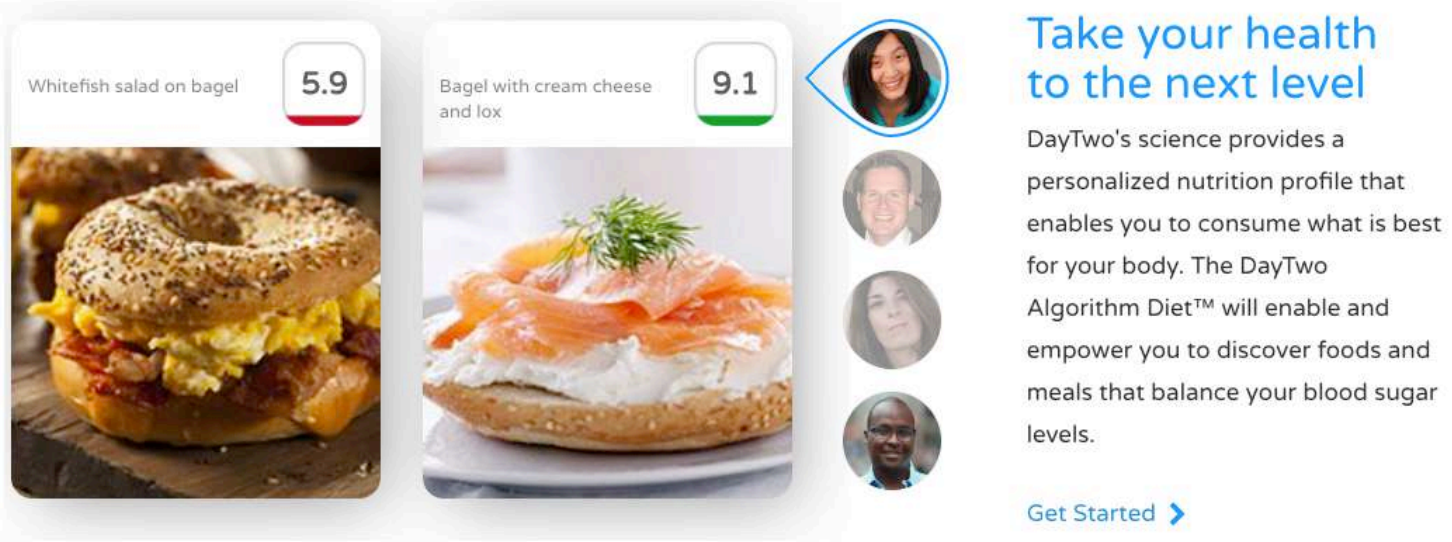
Fish density

Money Saved

Time spent to perform the experiment (before manually held)

Prescriptive Intelligence - DayTwo

- Based on readings of the user's gut microbiome, as well as the user's most recent biometrics, DayTwo predicts what food will cause more of a spike in the user's glucose levels. This is particularly useful for people suffering diabetes, or pre-diabetes.
- https://www.daytwo.com/en/



The Machine Learning Canvas (v0.4)

Designed for: DAYTWO | Designed by: | Date: | Iteration:

Decisions

How are predictions used to make decisions that provide the proposed value to the end-user?

The customer can modify its diet based on a personalized scheme of health. The customer can avoid choices which are known to be unhealthy at the individual level.

ML task

Input, output to predict, type of problem.

Input: User biometrics and gut microbiome sample. Particular meal combination.

Outputs: Rating of the meal according to how beneficial it is to the customer.

Value Propositions

What are we trying to do for the end-user(s) of the predictive system? What objectives are we serving?

What? A personalized rating system of meals, based on their adequacy for the user.

Who? All people with a particular interest in decreasing their spikes in glucose levels after a meal. Particularly, people suffering from diabetes or obesity.

Data Sources

Which raw data sources can we use (internal and external)?

Initially, data sources may be external, extracted from known test cases.
As a network of customers is developed, internal data sources may be obtained from them.

Collecting Data

How do we get new data to learn from (inputs and outputs)?

Measure the glucose levels of a subset of customers, and keep record of the meals they ingest. Based on this, keep improving the model, taking into account their gut microbiome types.

Making Predictions

When do we make predictions on new inputs? How long do we have to featurize a new input and make a prediction?

Make a prediction every time a new user-meal combination is proposed. Ideally, a ready-made list of meals is evaluated for each customer, such that a set of predictions is made by default for every new customer. Since only one prediction is made for each new sample, there is a relatively large amount of available time.

Offline Evaluation

Methods and metrics to evaluate the system before deployment.

Before deployment, employ a reduced test case of users, and measure their glucose levels in order to measure the accuracy of the model with different meals.

Features

Input representations extracted from raw data sources.

Gut microbiome.
Recent user biometrics.
Recently ingested meal.
Evolution of glucose levels.

Building Models

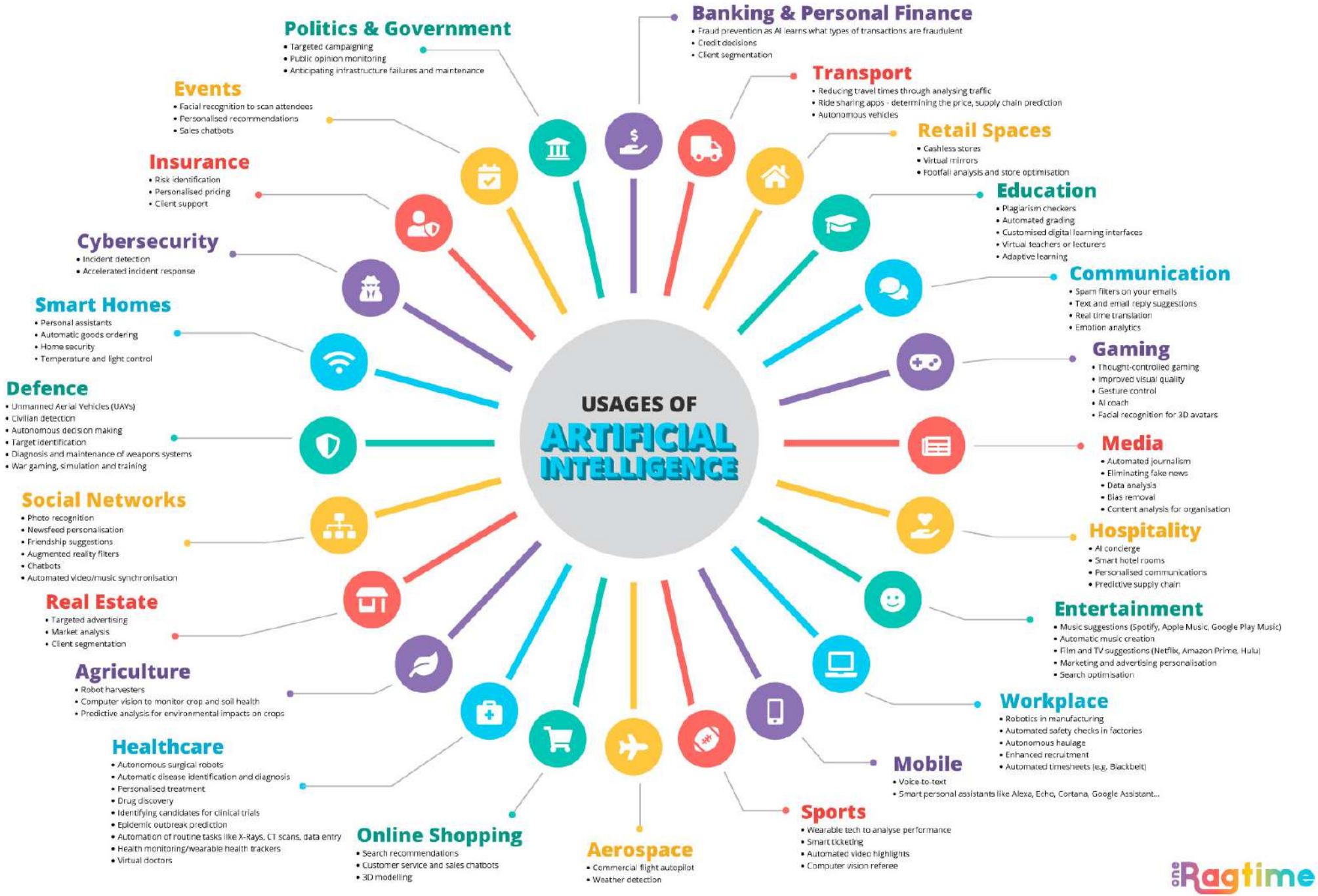
When do we create/update models with new training data? How long do we have to featurize training inputs and create a model?

Create a model for each gut microbiome type. Develop them based on new data compiled from customers supplying glucose level data.

Live Evaluation and Monitoring

Methods and metrics to evaluate the system after deployment, and to quantify value creation.

Value creation: Observe which meals are usually chosen by customers. To what extent are recommendations followed?
Measure model accuracy with glucose level data.


USAGES OF
ARTIFICIAL INTELLIGENCE
Politics & Government
• Targeted campaigning
• Public opinion monitoring
• Anticipating infrastructure failures and maintenance
Banking & Personal Finance
• Fraud prevention as AI learns what types of transactions are fraudulent
• Credit decisions
• Client segmentation
Events
• Facial recognition to scan attendees
• Personalised recommendations
• Sales chatbots
Transport
• Reducing travel times through analysing traffic
• Ride sharing apps - determining the price, supply chain prediction
• Autonomous vehicles
Insurance
• Risk identification
• Personalised pricing
• Client support
Retail Spaces
• Cashless stores
• Virtual mirrors
• Footfall analysis and store optimisation
Cybersecurity
• Incident detection
• Accelerated incident response
Education
• Plagiarism checkers
• Automated grading
• Customised digital learning interfaces
• Virtual teachers or lecturers
• Adaptive learning
Smart Homes
• Personal assistants
• Automatic goods ordering
• Home security
• Temperature and light control
Communication
• Spam filters on your emails
• Text and email reply suggestions
• Real time translation
• Emotion analytics
Defence
• Unmanned Aerial Vehicles (UAVs)
• Civilian detection
• Autonomous decision making
• Target identification
• Diagnosis and maintenance of weapons systems
• War gaming, simulation and training
Gaming
• Thought-controlled gaming
• Improved visual quality
• Gesture control
• AI coach
• Facial recognition for 3D avatars
Social Networks
• Photo recognition
• Newsfeed personalisation
• Friendship suggestions
• Augmented reality filters
• Chatbots
• Automated video/music synchronisation
Media
• Automated journalism
• Eliminating fake news
• Data analysis
• Bias removal
• Content analysis for organisation
Real Estate
• Targeted advertising
• Market analysis
• Client segmentation
Hospitality
• AI concierge
• Smart hotel rooms
• Personalised communications
• Predictive supply chain
Agriculture
• Robot harvesters
• Computer vision to monitor crop and soil health
• Predictive analysis for environmental impacts on crops
Entertainment
• Music suggestions (Spotify, Apple Music, Google Play Music)
• Automatic music creation
• Film and TV suggestions (Netflix, Amazon Prime, Hulu)
• Marketing and advertising personalisation
• Search optimisation
Workplace
• Robotics in manufacturing
• Automated safety checks in factories
• Autonomous haulage
• Enhanced recruitment
• Automated timesheets (e.g. Blackbelt)
Healthcare
• Autonomous surgical robots
• Automatic disease identification and diagnosis
• Personalised treatment
• Drug discovery
• Identifying candidates for clinical trials
• Epidemic outbreak prediction
• Automation of routine tasks like X-Rays, CT scans, data entry
• Health monitoring/wearable health trackers
• Virtual doctors
Mobile
• Voice-to-text
• Smart personal assistants like Alexa, Echo, Cortana, Google Assistant...
Online Shopping
• Search recommendations
• Customer service and sales chatbots
• 3D modelling
Aerospace
• Commercial flight autopilot
• Weather detection
Sports
• Wearable tech to analyse performance
• Smart ticketing
• Automated video highlights
• Computer vision referee