The Practical Machine Learning Series

Ensemble and Deep Learning

Why are you here?

On a scale of one to ten, how motivated are you to continue your machine learning, learning?

Why wasn't that number lower?

“Machine Learning is a toolkit of algorithms that finds insight from data”

—Matt Kirk

“Use linear regression, and when that fails try something else”

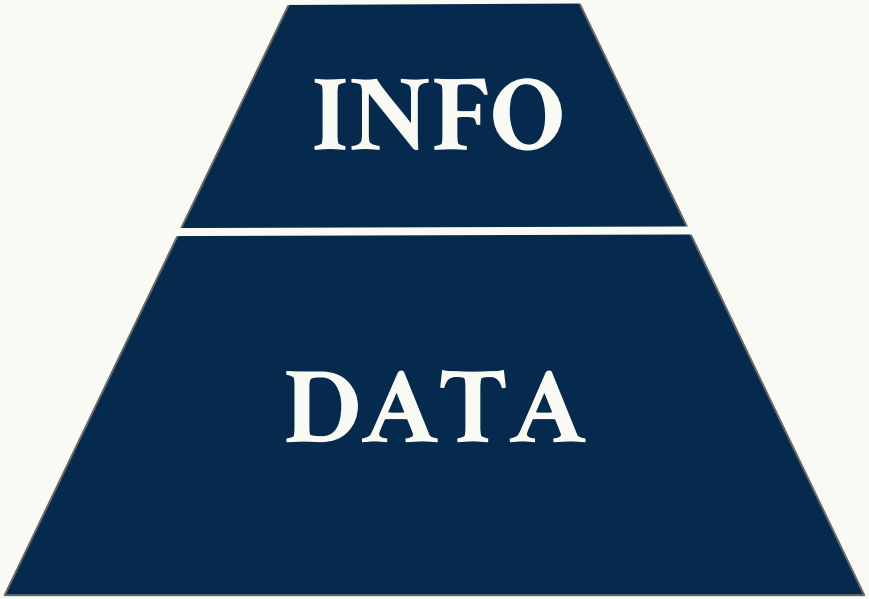
—Kevin Jamieson (UW Prof)

What is Machine Learning? DATA

- Temperatures
- Voltage
- Pressure
- Force
- Binary Data
- DOM events

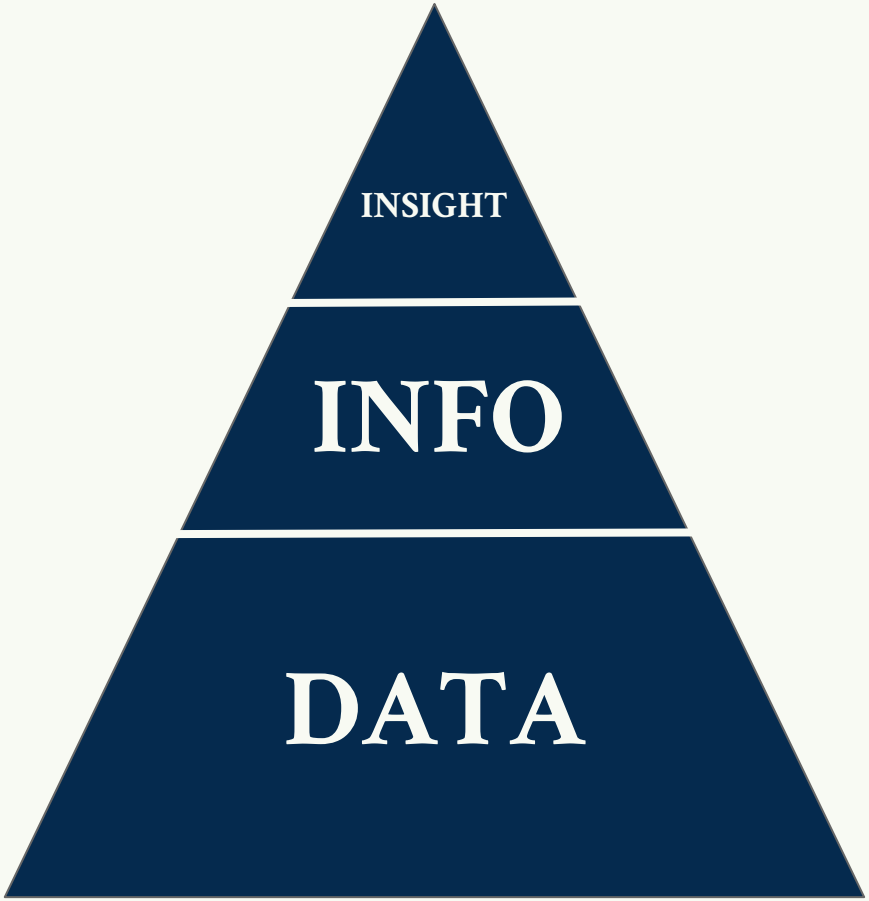
What is Machine Learning? INFO

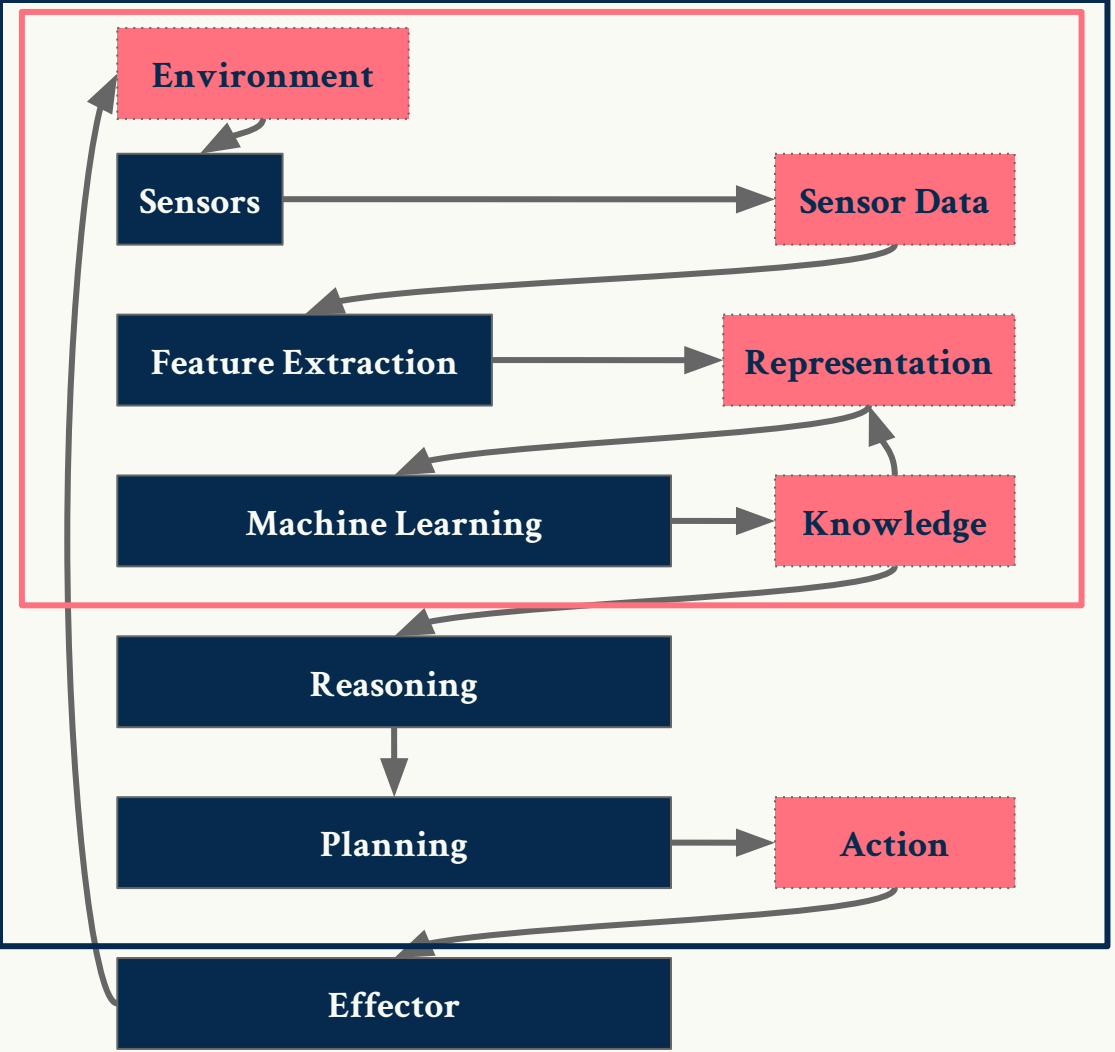
- Median
- Mode
- Average
- Kurtosis
- Skew
- The Record High
- The Record Low



What is Machine Learning? INSIGHT

- What is the predicted outcome?
- What is in the data?
- What is the best strategy?
- What is something surprising?





The Promise of Machine Learning

The Promise of Reinforcement Learning

“magic is pretty simple: It comes down to training, practice, and experimentation”

—David Blaine

Our Work Today

1

The four directions

Feature learning, ensemble learning, deep learning. What to do when things go wrong

2

Ensemble Learning

Factorization, as well as ensemble learning to improve models substantially.

3

Shallow intro to Deep Learning

We will dip into deep learning to see how we can improve our situation even further

How you learn

LECTURE
Write notes during this time. Make sure you have questions!

QUIZ
Some simple questions to test your recall.

PRACTICE
I will share my screen and walk through the practice of the work.

TRY ON
This is where you try machine learning on for yourself! Exercise!

Who am I speaking with today?

CTO

Wants to learn how to use Machine Learning for her business.

Analyst

Wants to level up their ability to analyse data.

Developer

Wants to understand what data scientists do all day.

Learner

Just is continuing their learning path!

Matt Kirk

matt@foreshadow.ai
matt@yourchiefscientist.com
matt@matthewkirk.com
DataScienceSuperHero.com

1. Four Directions to improve your model

- Data Engineering
- Feature Engineering
- Ensemble Learning
- Deep Learning
- How do we know we've improved?

“What algorithm should I pick? How do I improve my ML model?”

—FAQ

The Three Classes of Machine Learning

SUPERVISED

$$f(x) = y$$

Learn a function that takes an input and outputs a class or value.

UNSUPERVISED

$$f(x) = x$$

Learn how to construct the input data with less complexity.

REINFORCEMENT

$$f_t(x_t) = y_{t+1}$$

Learn how to predict the next outcome over time.

The Three Classes of Machine Learning

SUPERVISED

$$f(x) = y$$

Learn a function that takes an input and outputs a class or value.

UNSUPERVISED

$$f(x) = x$$

Learn how to construct the input data with less complexity.

REINFORCEMENT

$$f_t(x_t) = y_{t+1}$$

Learn how to predict the next outcome over time.

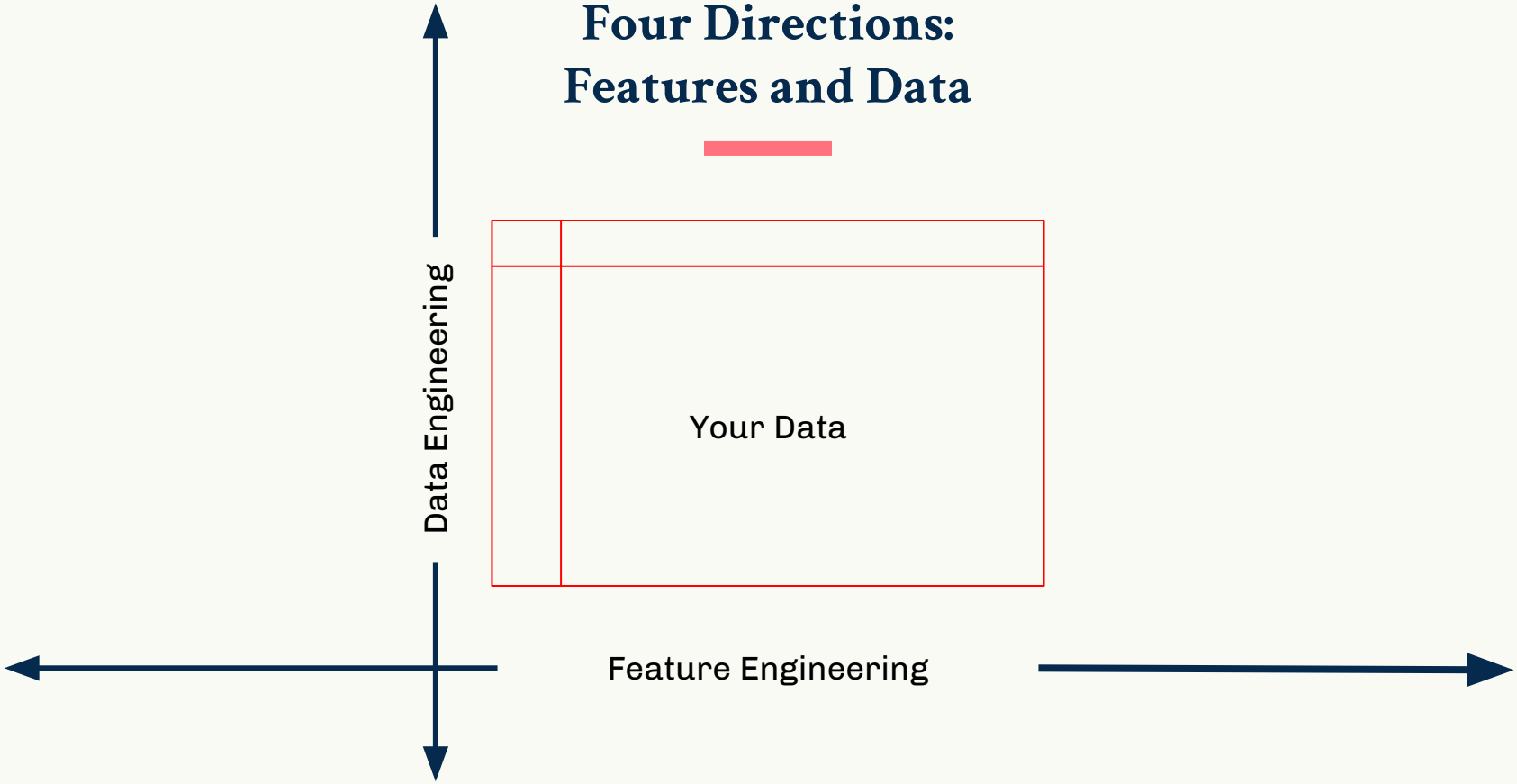
Composing Unsupervised with Supervised Learning

$$f(x) = y$$

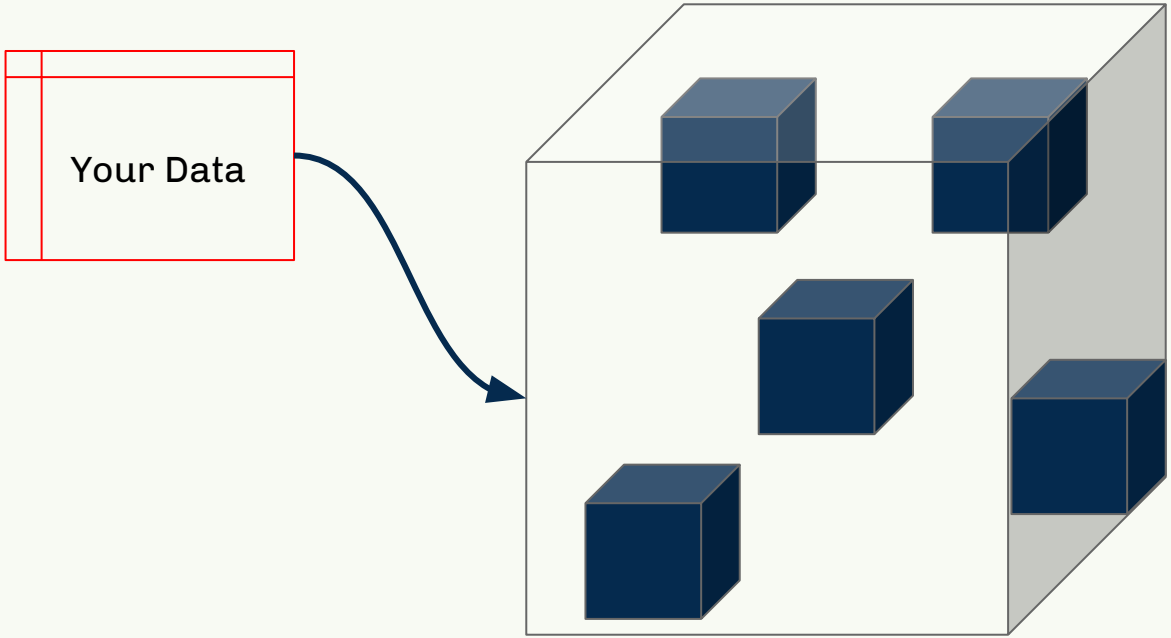
$$g(x) = x$$

$$\therefore f(g(x)) = y$$

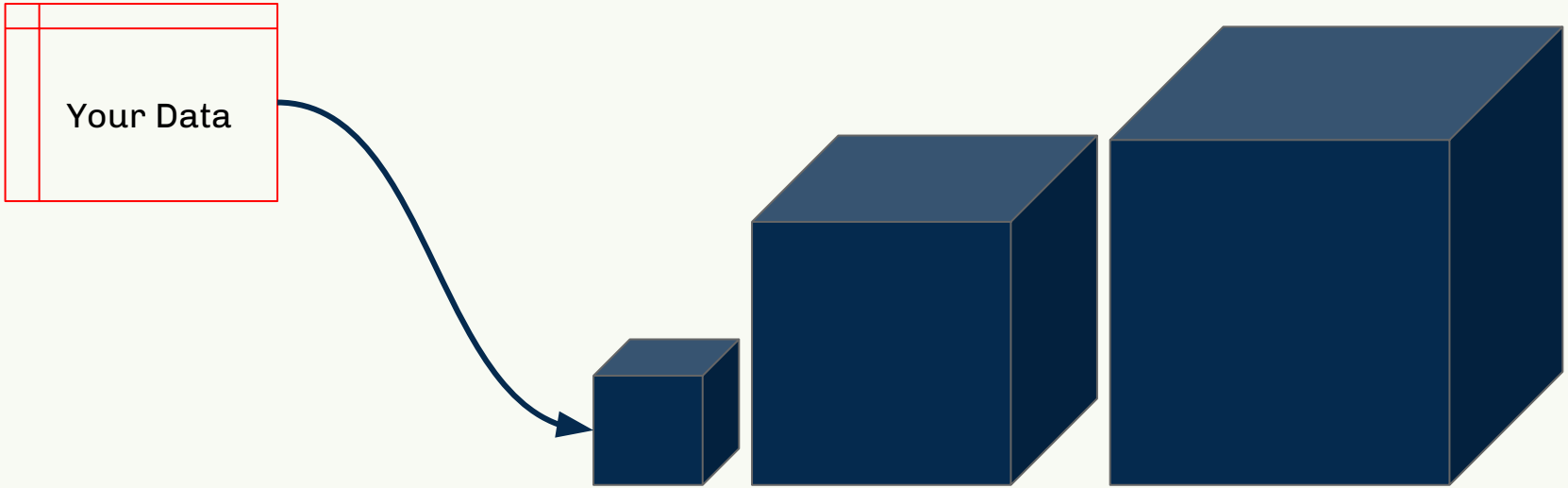
Four Directions: Features and Data



Four Directions: Ensemble



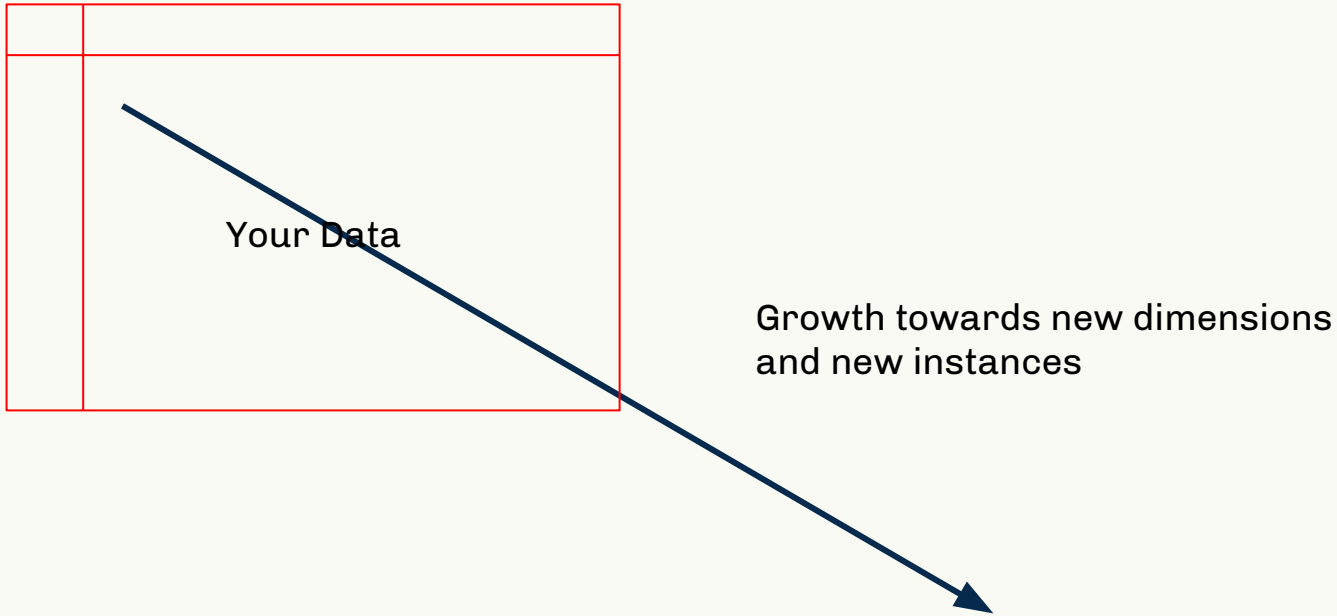
Four Directions:
Deep Learning



Data Engineering

Move southeast

Data Engineering goal



Feature Engineering

Using what you got the best way possible

Clarity:
Dimension vs Feature

<51.178882, -1.826215>

Clarity:
Dimension vs Feature

-1

“Terrible!”

Feature Engineering

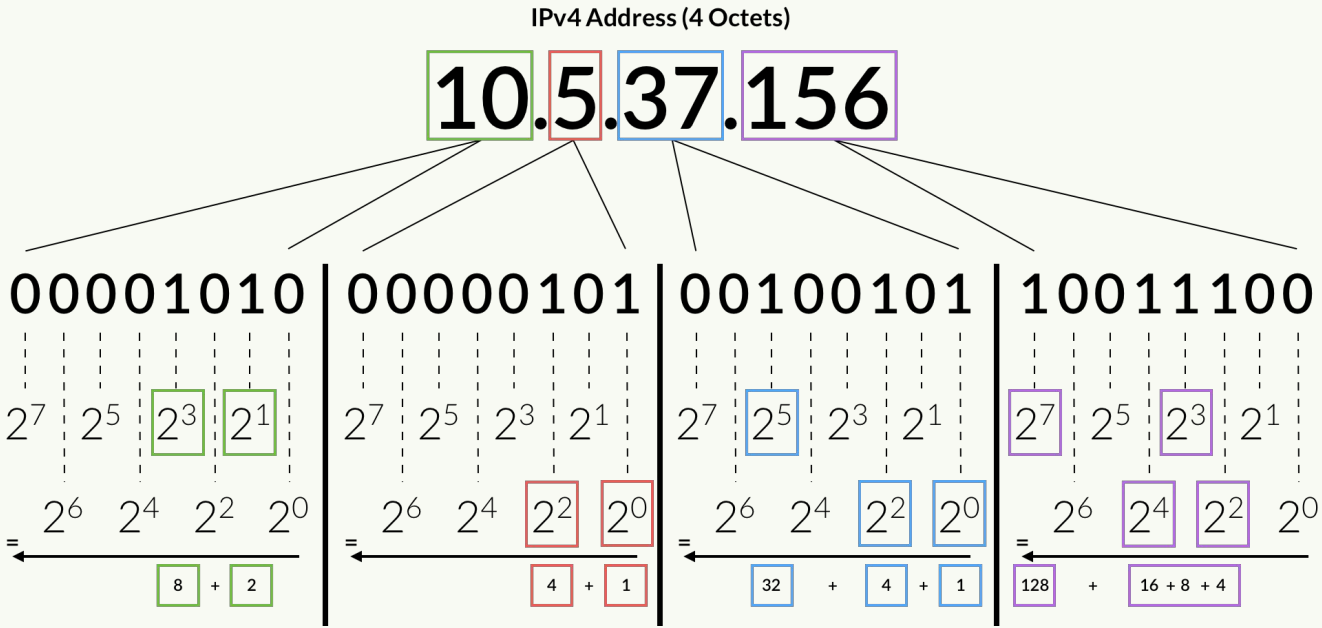
RULES
IPV4 is an integer.
Latitude and longitude can be transformed into geohash.
Domain specific rules.

SELECTION
Select the best dimension out of the dataset

TRANSFORM
Yes, this is the ringed one. It’s a gas giant, composed mostly of hydrogen and helium

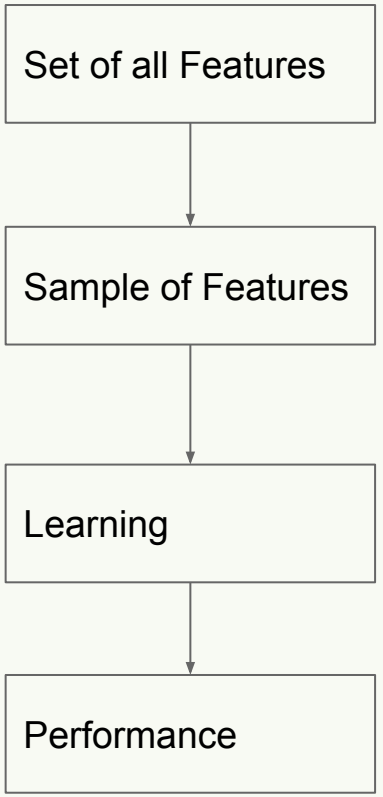
LEARN
Neptune is the farthest planet in our Solar System and also the fourth-largest

Rules: IPv4 is just a number



Feature Selection

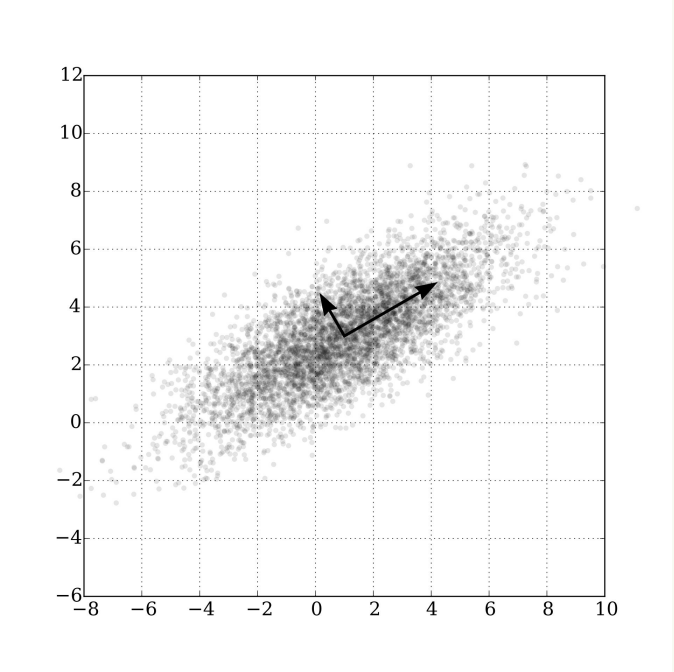
- Random Selection
- Tree Based
- L1-based



https://scikit-learn.org/stable/modules/feature_selection.html

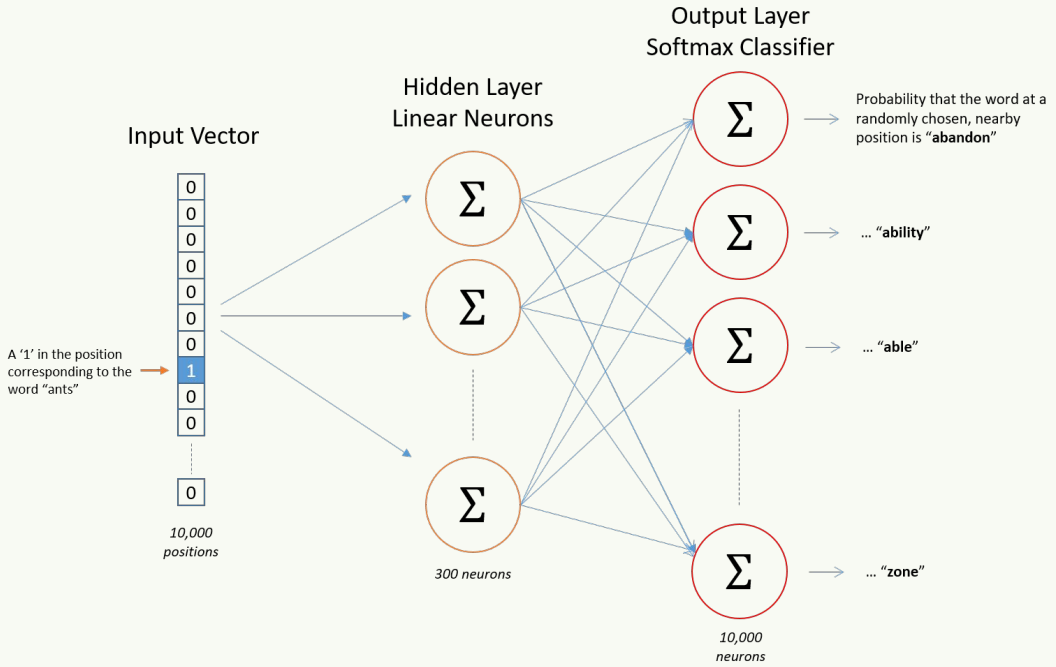
Feature Transform

- PCA
- ICA
- tSNE
- UMAP



Feature Learning

- word2vec
- Autoencoder
- Embeddings



Ensemble Learning

Weather Patterns

Monte Carlo Generators

Ensemble Learning

RULES
Weather models, risk models.

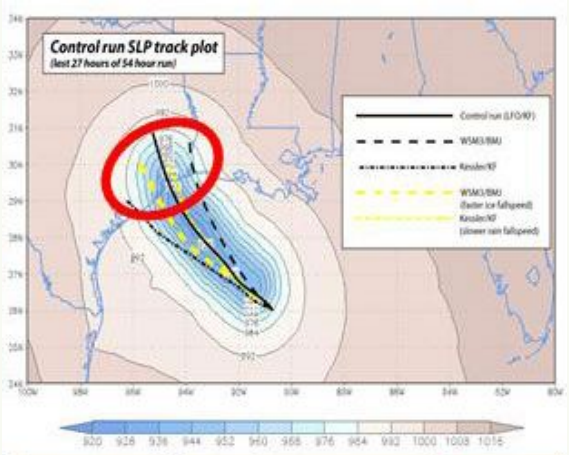
BAGGING
Building many models together and averaging or finding the median of answers.

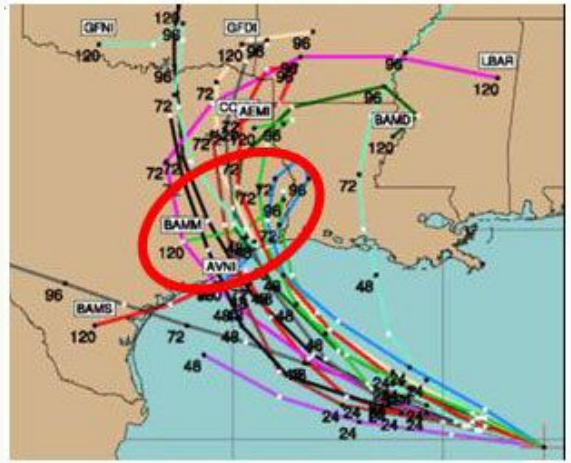
BOOSTING
Combining multiple models together via a weighted and learned mechanism

SELECTING
Hyper parameter tuning is a form of ensemble learning. Selecting the best performing one.

Rules

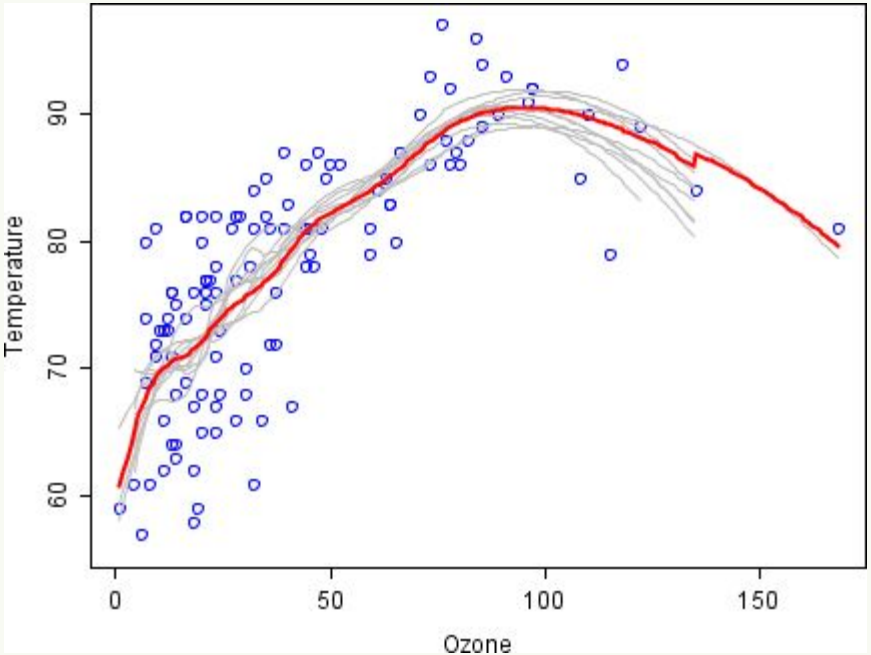
- Weather Models
- Risk Models
- Actuarial Tables





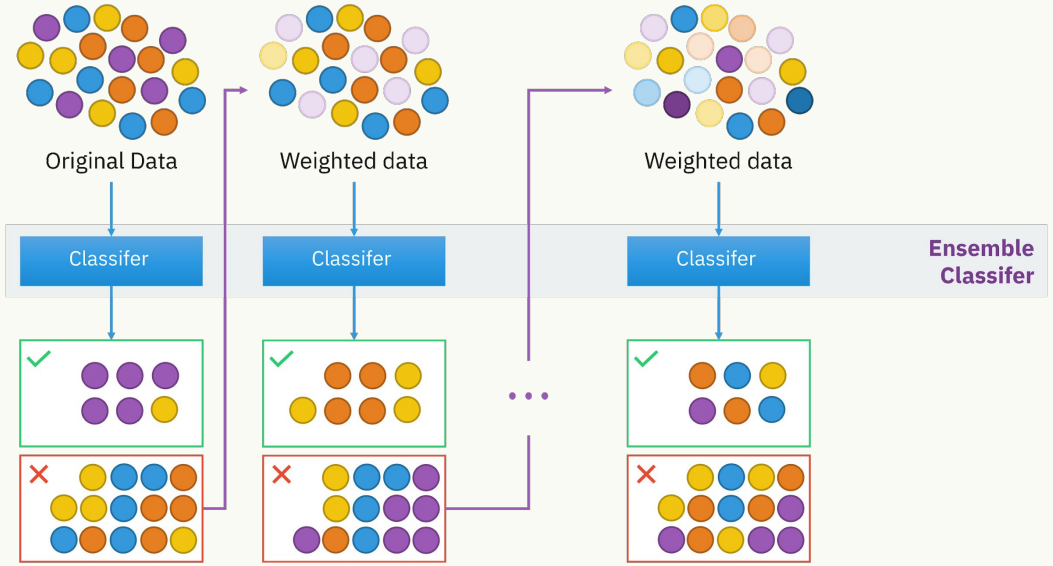
Bagging

- Simple model introduced by Leo Breiman
- Learn a bunch of models and average them!



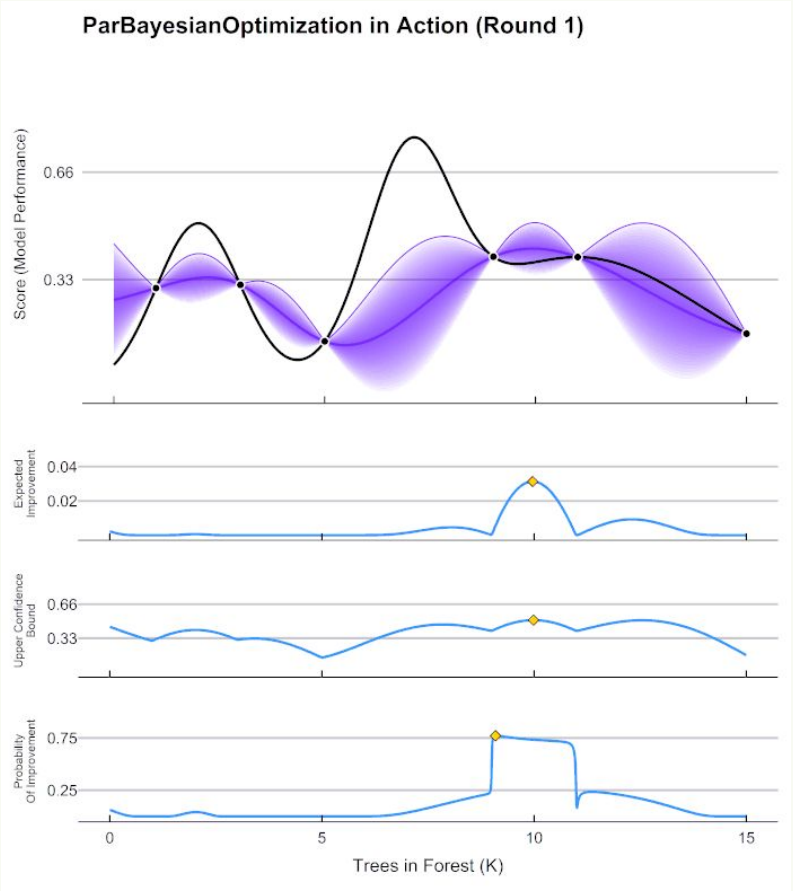
Boosting

- AdaBoost
- CatBoost
- XgBoost
- GBM
- EBM
- **Wins Kaggle Competitions**



Model Selection

- Grid Search
- Gaussian Process
- Newtons Method
- Bayesopt
- LIPO
- **Global Optimizations (tough to do in practice)**



Deep Learning

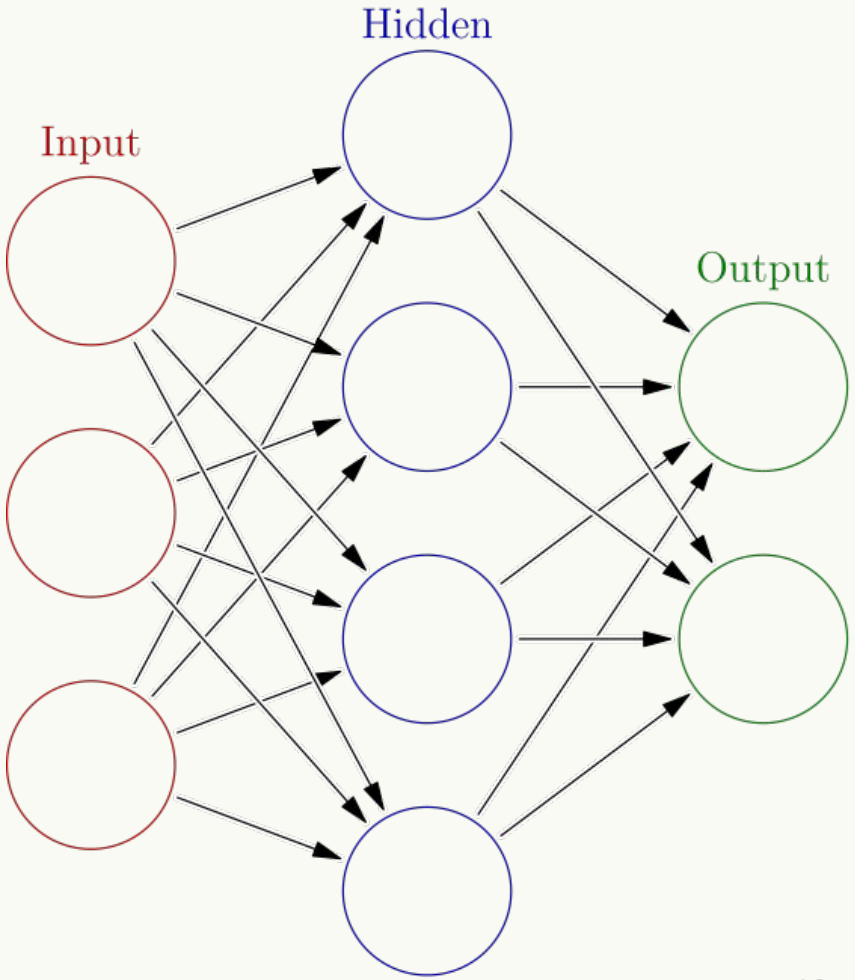
Just keep adding neurons

What is Deep Learning?

“Feature Learning”
~ Yoshua Bengio

“Large Neural Nets”
~ Andrew Ng

ANNs



Bottleneck

- Theory as to why deep learning works in the first place at all

The Four Directions

MORE/LESS DATA

Adding data will always add statistical power to the models. Subtracting might help with outliers.

MORE/LESS DIMENSIONS

More dimensions sometimes finds a new method. Sometimes they are irrelevant.

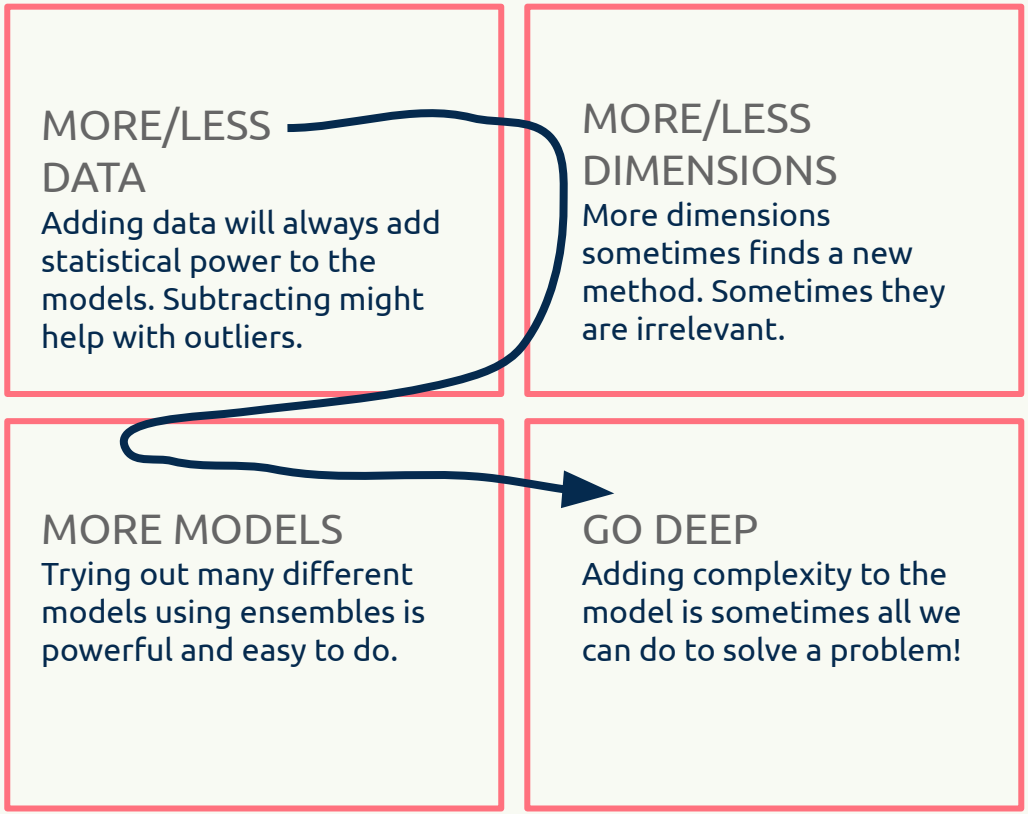
MORE MODELS

Trying out many different models using ensembles is powerful and easy to do.

GO DEEP

Adding complexity to the model is sometimes all we can do to solve a problem!

The Four Directions



QUIZ #1!

What is a form of feature engineering?

1. Turning IP addresses into numbers.
2. Learning a vector for a word.
3. Turning hundreds of dimensions into a select few.
4. All of the above

QUIZ #1!

What is a form of feature engineering?

1. Turning IP addresses into numbers.
2. Learning a vector for a word.
3. Turning hundreds of dimensions into a select few.
4. All of the above

QUIZ #1!

How can unsupervised and supervised learning be used together?

1. To compare and contrast the approaches
2. To compose unsupervised inside of supervised
3. To compose supervised inside of unsupervised
4. To prevent your models from growing exponentially

QUIZ #1!

How can unsupervised and supervised learning be used together?

1. To compare and contrast the approaches
2. To compose unsupervised inside of supervised
3. To compose supervised inside of unsupervised
4. To prevent your models from growing exponentially

QUIZ #1!

When should you use ensemble learning over deep learning?

1. When you don't have a huge budget
2. When you don't have a lot of time
3. When the precision doesn't need to be exceptionally high
4. All of the above

QUIZ #1!

When should you use ensemble learning over deep learning?

1. When you don't have a huge budget
2. When you don't have a lot of time
3. When the precision doesn't need to be exceptionally high
4. All of the above

Break time

Next up is UMAP and Random Forests

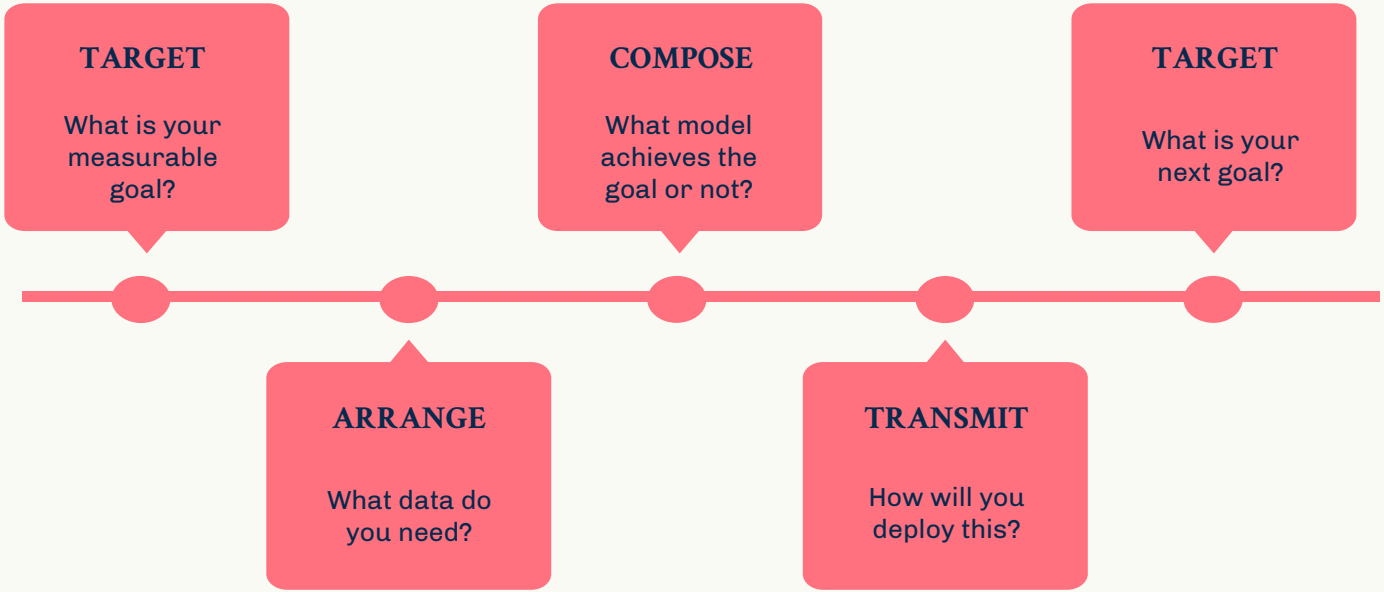
2. UMAP and CatBoost

- What is matrix Factorization?
- What is ensemble learning?
- Why are they powerful?
- Feature Engineering
- Bagging
- Boosting
- XgBoost / CatBoost

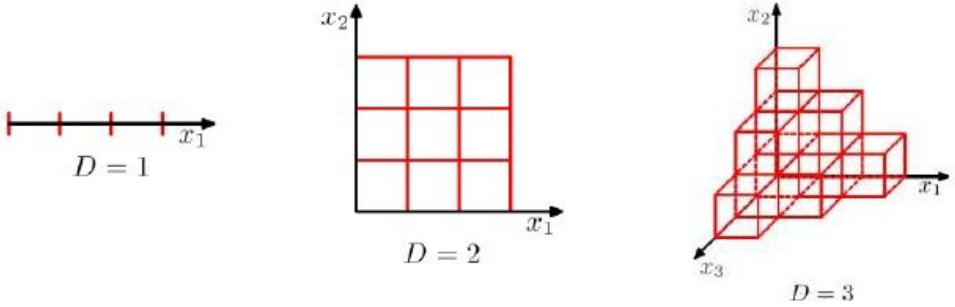
“Use linear regression, and when that doesn’t work try something else”

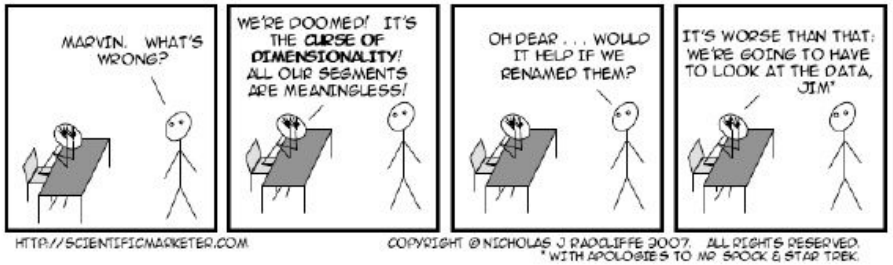
—Kevin Jamieson

T.A.C.T.


TARGET
What is your measurable goal?
ARRANGE
What data do you need?
COMPOSE
What model achieves the goal or not?
TRANSMIT
How will you deploy this?
TARGET
What is your next goal?

Less is More
The Curse of Dimensionality
(Bellman, 1961)





Feature Engineering

Using what you got the best way possible

Feature Engineering

RULES
IPV4 is an integer.
Latitude and longitude can be transformed into geohash.
Domain specific rules.

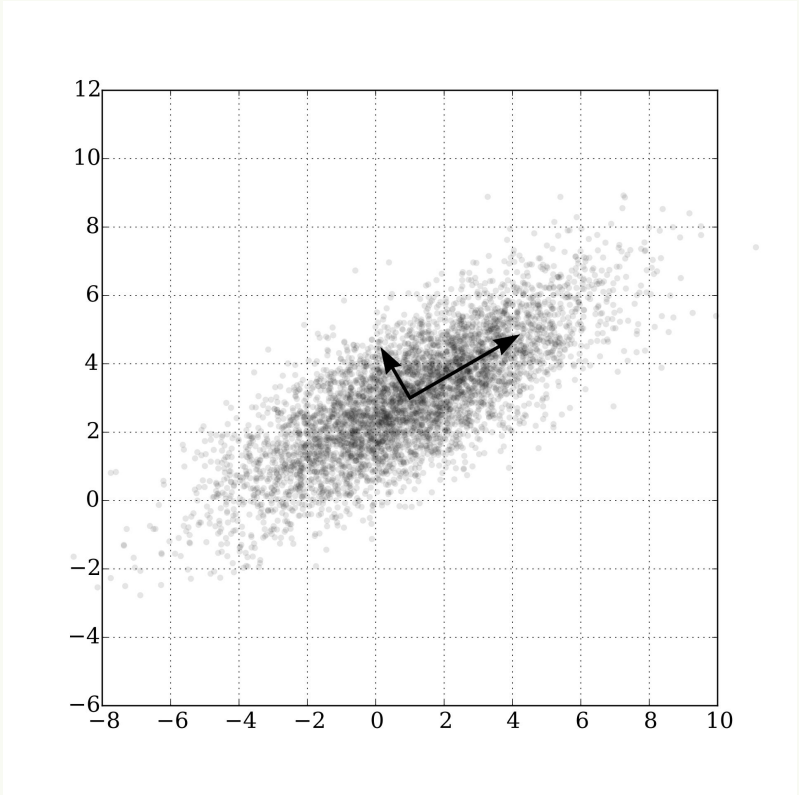
SELECTION
Select the best dimension out of the dataset

TRANSFORM
Yes, this is the ringed one. It’s a gas giant, composed mostly of hydrogen and helium

LEARN
Neptune is the farthest planet in our Solar System and also the fourth-largest

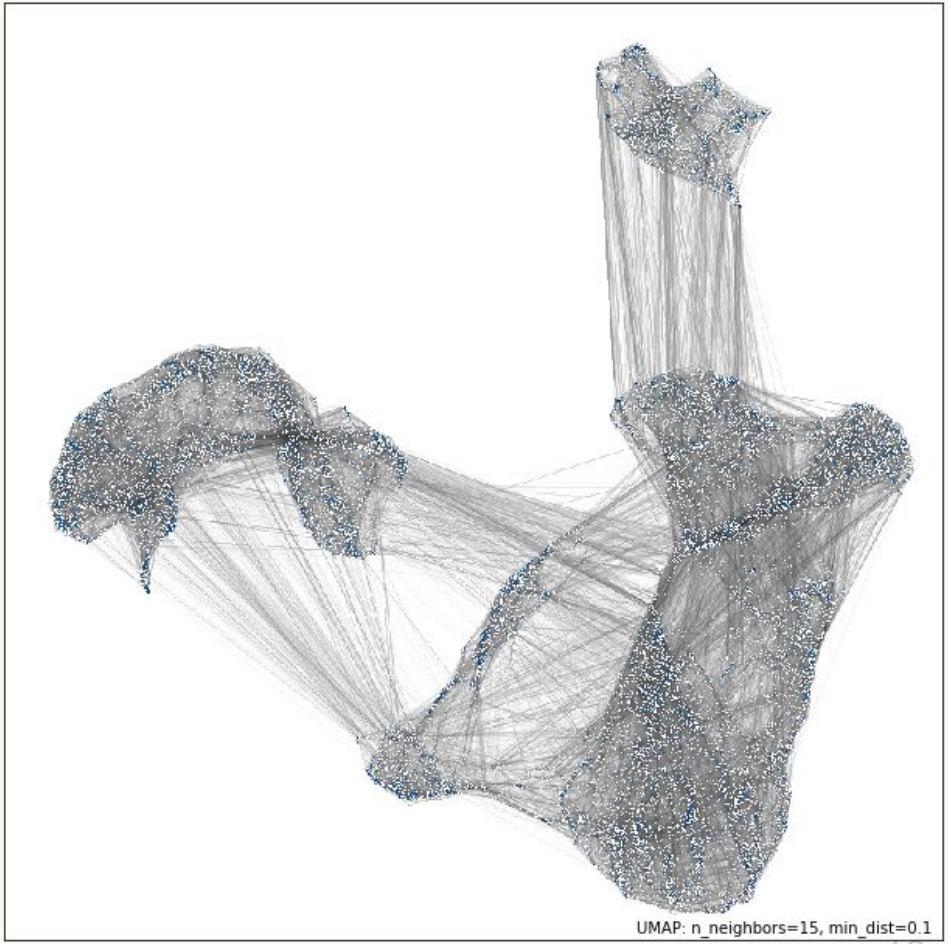
PCA

- Linear
- Simple
- Components



UMAP

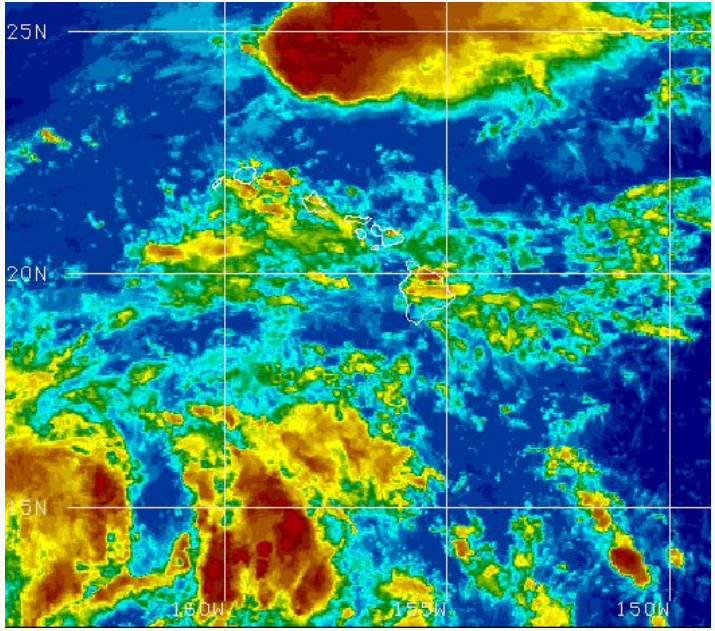
- Manifold
- Nearest Neighbors
- Faster than tSNE
- Works very well!
- https://towardsdatascience.com/how-exactly-umap-works-13e3040e1668



Ensemble Learning

Nothing says we can’t use more than one model

Weather Forecasting

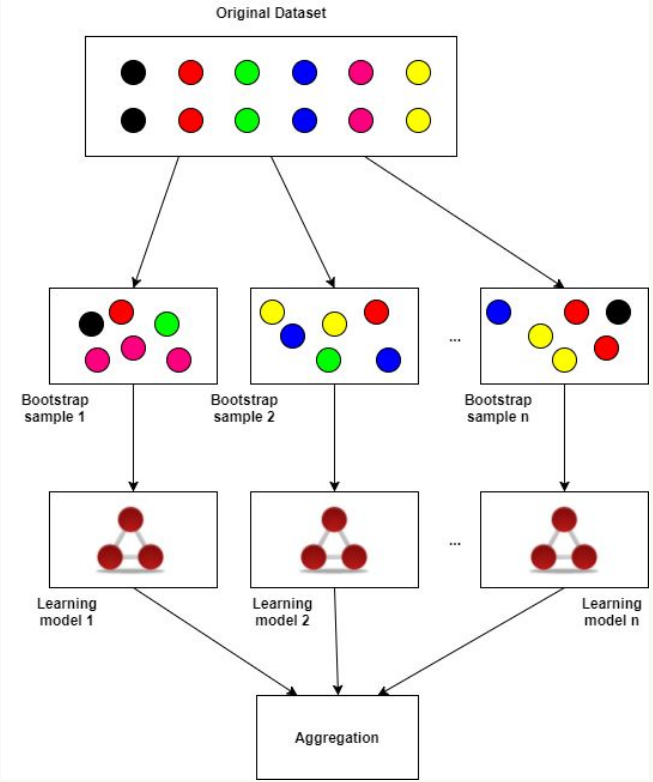


Overfit vs Underfit

Bagging

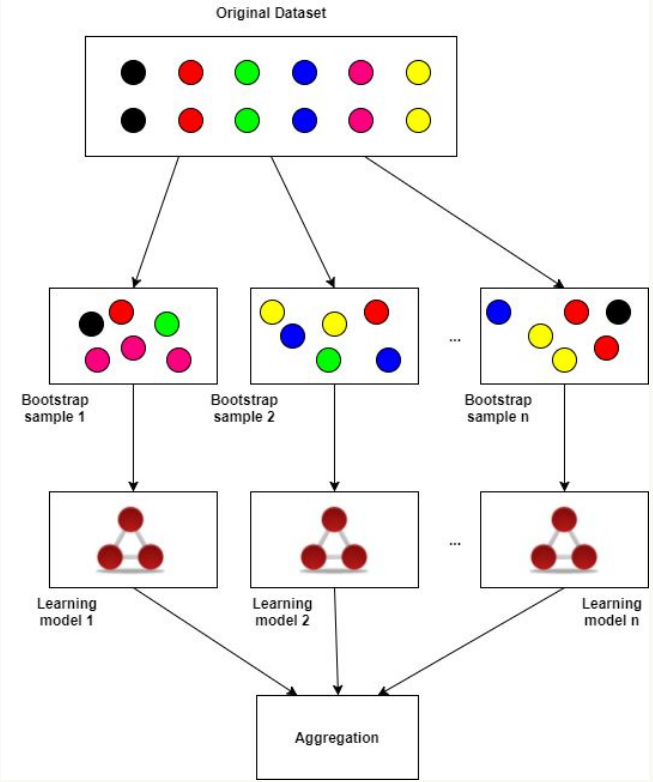
Sample N times and train N models and then just average the results together!

Works on all supervised learning models.



Boosting

Same thing but you will learn the optimal weighting based on tests iteratively



“Watch what gets used on Kaggle (CatBoost, XgBoost, etc)”

QUIZ #2!

Why is UMAP more powerful than PCA?

1. PCA is purely linear
2. UMAP was invented by smarter mathematicians
3. UMAP is faster
4. PCA is too much math

QUIZ #2!

Why is UMAP more powerful than PCA?

1. PCA is purely linear
2. UMAP was invented by smarter mathematicians
3. UMAP is faster
4. PCA is too much math

QUIZ #2!

What is bagging?

1. Giving up on a machine learning model that won't work
2. Grouping many models together and averaging them
3. A new optimization technique from groceries
4. Not useful

QUIZ #2!

What is bagging?

1. Giving up on a machine learning model that won't work
2. Grouping many models together and averaging them
3. A new optimization technique from groceries
4. Not useful

QUIZ #2!

What is boosting?

1. Adding hardware to your ML models
2. Powerups for your ML models
3. Aggregating many models together in a weighted fashion
4. Nootropics for your models

QUIZ #2!

What is boosting?

1. Adding hardware to your ML models
2. Powerups for your ML models
3. Aggregating many models together in a weighted fashion
4. Nootropics for your models

Your turn

Demo and then it's your turn!!

Break time

Next up is a teensy step into deep learning!

3. Bidirectional LSTM

- Why Neural Nets are so powerful.
- Representation Learning
- Recurrent Neural Networks
- The cost of DL

Neural Nets

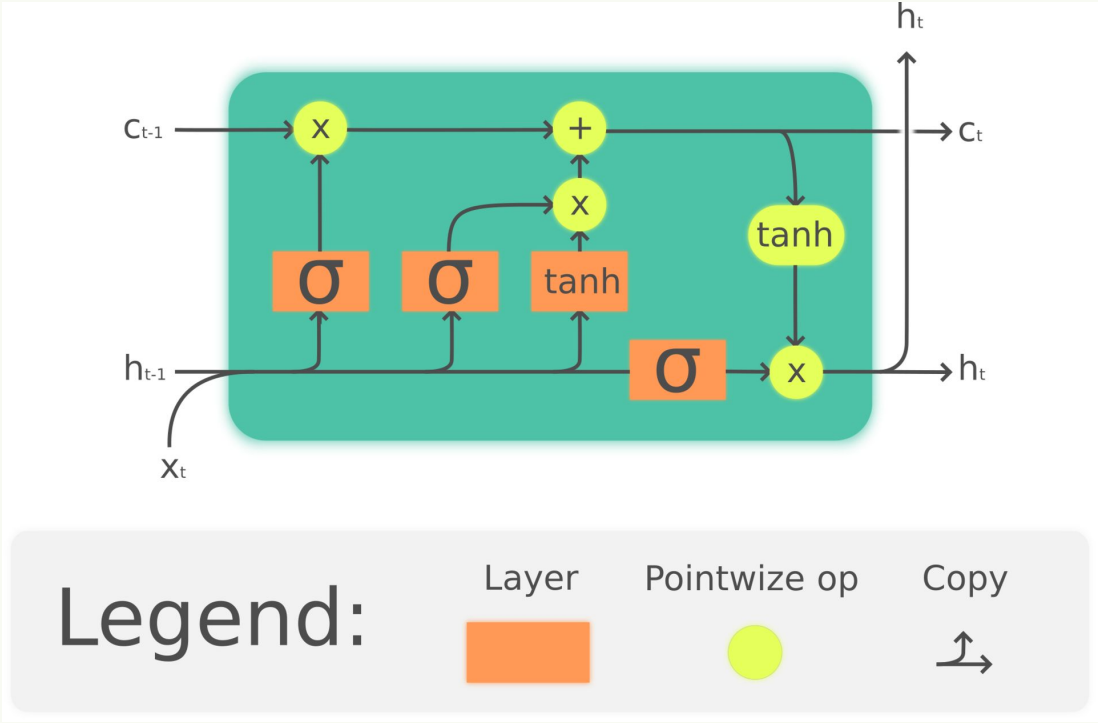
The original research on neural networks was to try and recreate the brain 1:1. Biological focused.

7 +/- 2

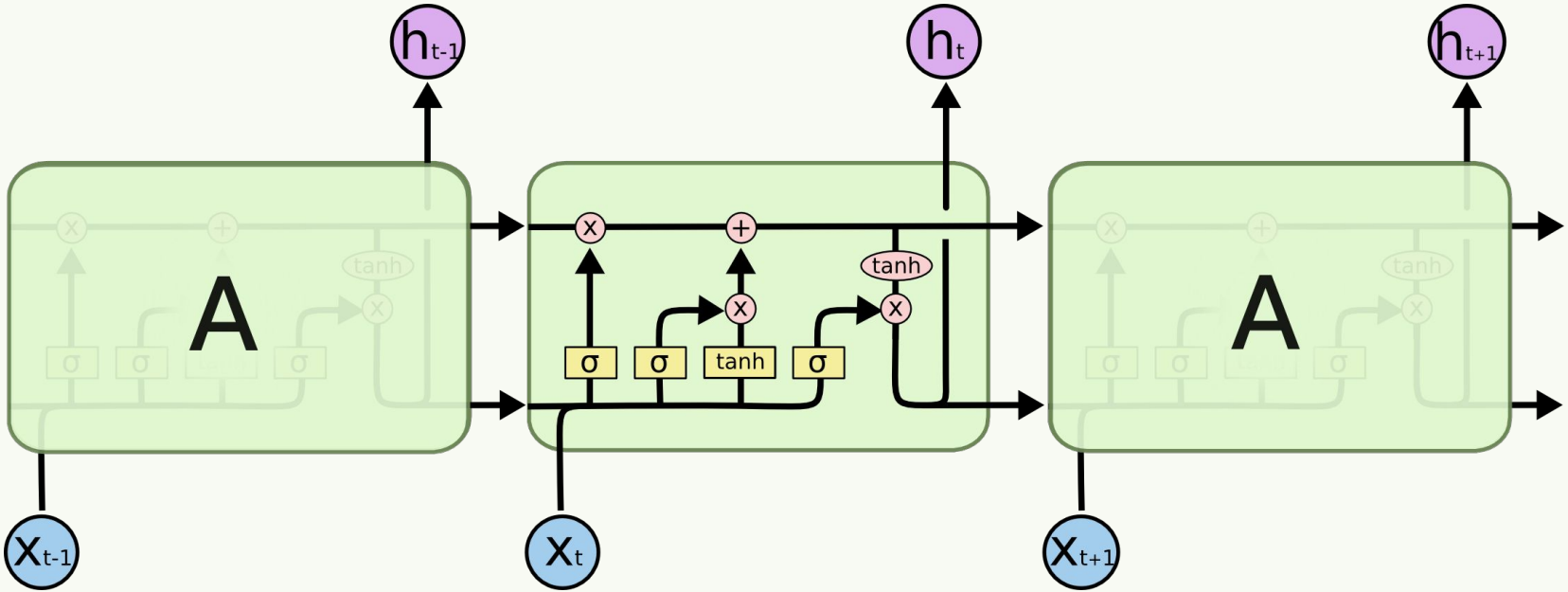
- Working memory hypothesis
- Phonological loop
- Visual spatial sketchpad
- Central Executive

Reality

- Still not quite there (though we'll see with GPT-3)

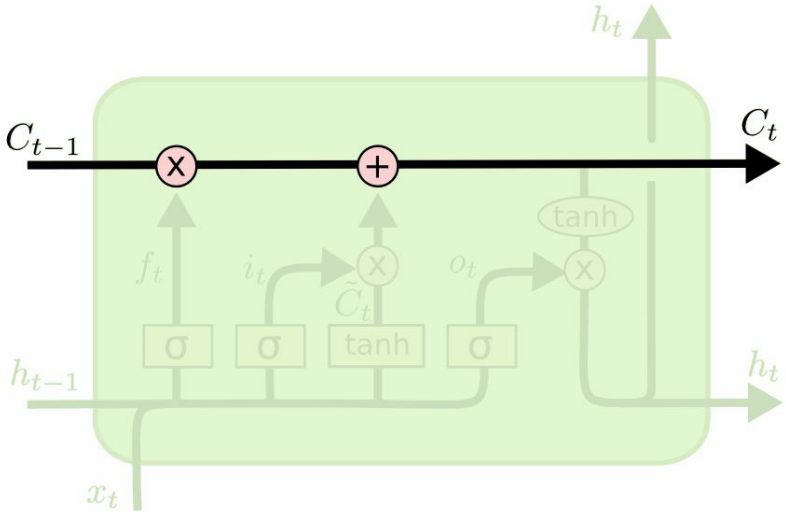


LSTM: Long Short Term Memory



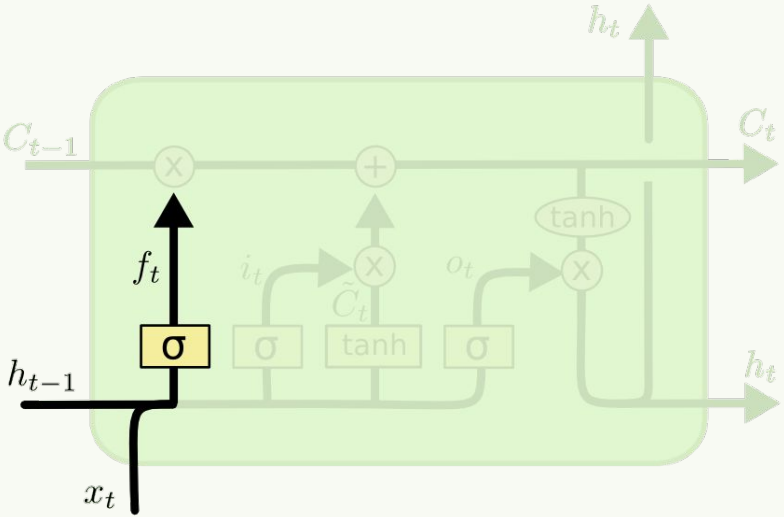
https://colah.github.io/posts/2015-08-Understanding-LSTMs/

LSTM: Long Short Term Memory
Keep a working memory



https://colah.github.io/posts/2015-08-Understanding-LSTMs/

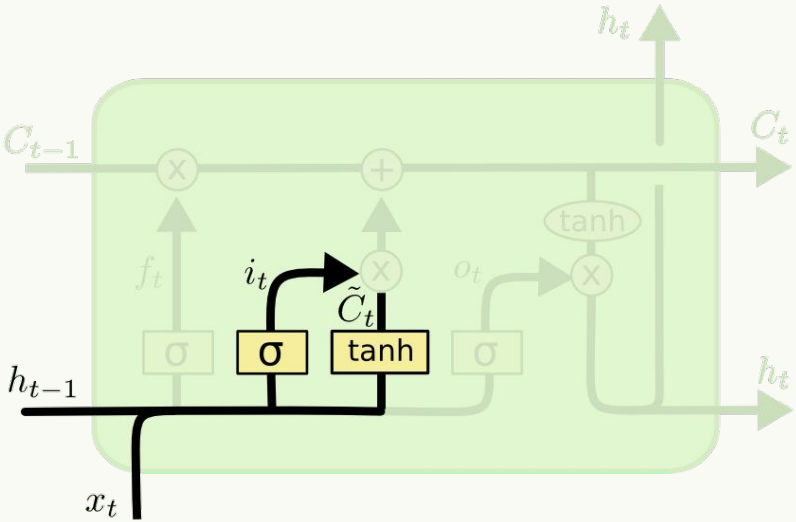
LSTM: Long Short Term Memory
Throw away old state



$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] \; + \; b_f\right)$$

https://colah.github.io/posts/2015-08-Understanding-LSTMs/

LSTM: Long Short Term Memory
Learn new state to retain

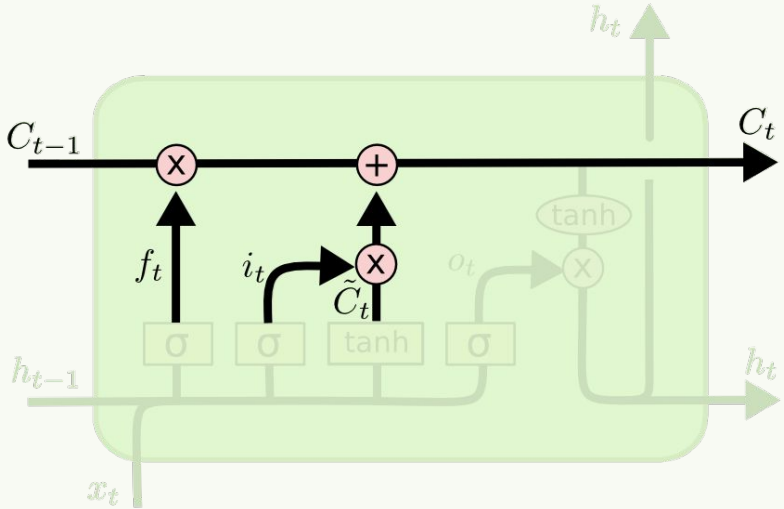


$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] \; + \; b_i\right)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] \; + \; b_C)$$

https://colah.github.io/posts/2015-08-Understanding-LSTMs/

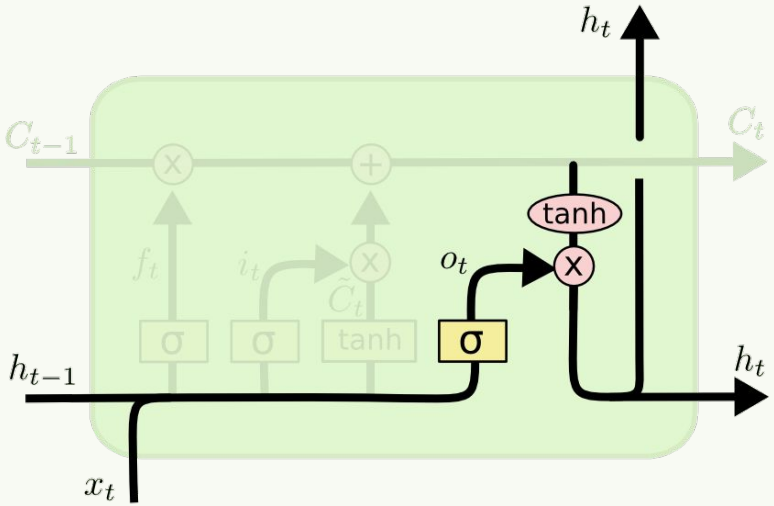
LSTM: Long Short Term Memory
Learn new state to retain



$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

https://colah.github.io/posts/2015-08-Understanding-LSTMs/

LSTM: Long Short Term Memory
What to output



$$o_t = \sigma\left(W_o\left[h_{t-1}, x_t\right] + b_o\right)$$

$$h_t = o_t * \tanh\left(C_t\right)$$

https://colah.github.io/posts/2015-08-Understanding-LSTMs/

Scratching the surface

- Convolutions
- Recurrent
- Transfer Learning
- Transformer networks
- GNNs
- Autoencoders
- Generative Models
- Probabalistic Models
- Representation Models
- Deep Reinforcement Learning
- Single Shot Object Segmentation
- More coming out all the time

QUIZ #3!

What does LSTM stand for?

1. Long Short Trained Model
2. Long Short Traffic Memory
3. Long Short Transient Model
4. Long Short Term Memory

QUIZ #3!

What does LSTM stand for?

1. Long Short Trained Model
2. Long Short Traffic Memory
3. Long Short Transient Model
4. Long Short Term Memory

QUIZ #3!

Why are LSTMs powerful?

1. They memorize a sequence
2. They operate fast
3. They learn what to store in a cell state and what to forget
4. Because medium told me they were

QUIZ #3!

Why are LSTMs powerful?

1. They memorize a sequence
2. They operate fast
3. They learn what to store in a cell state and what to forget
4. Because medium told me they were

QUIZ #3!

Why is it hard to re-create our brains?

1. They are incredibly complex
2. Our short term memory is incredibly small
3. Our long term memory is huge
4. All of the above

QUIZ #3!

Why is it hard to re-create our brains?

1. They are incredibly complex
2. Our short term memory is incredibly small
3. Our long term memory is huge
4. All of the above

Your turn

Demo and then it's your turn!!

We covered a lot today!

FOUR DIRECTIONS

Data, Dimensions, Models, and Depth

FEATURE ENGINEERING

Feature selection, transformations, rules, and representations

ENSEMBLE LEARNING

Rule based learning, bagging, boosting, and selection

DEEP LEARNING

Feature learning + big neural nets but also LSTMs

UMAP AND ENSEMBLES

Dimension reduction using UMAP and also ensemble learning with SK-LEARN

LSTMS

Long short term memory and recurrent neural networks!

What was your number?

What are you going to do now?

Three months from now

Six months from now

One year from now

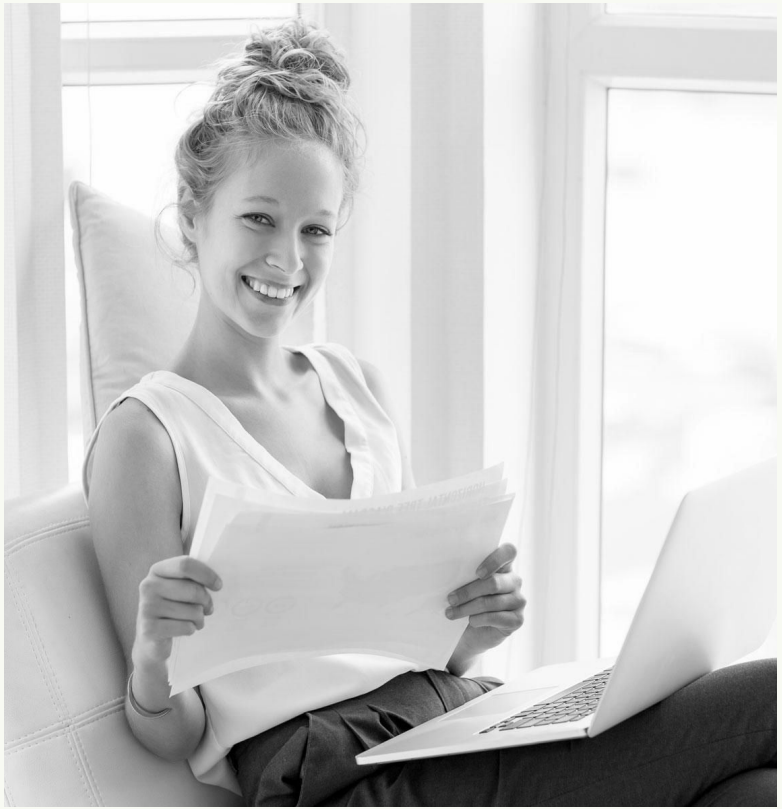
Thanks!

matt@matthewkirk.com

datasciencesuperhero.com
yourchiefscientist.com
foreshadow.ai

Whoa!

It could be the part of the presentation where you can introduce yourself, write your email...



Important information

Mercury’s name has nothing to do with the liquid metal since it was named after the Roman messenger god, Mercury

> **“This is a quote. Words full of wisdom that someone important said and can make the reader get inspired.”**
>
> —Someone famous

1. This is a great headline

And this is the subtitle that
makes it comprehensible

The slide title goes here!

Do you know what helps you make your point clear? Lists like this one:

Because they’re simple
You can organize your ideas clearly
And because you’ll never forget to buy milk!

And the most important thing: the audience won’t miss the point of your presentation

Maybe you need to divide the content

MERCURY

Mercury is the closest planet to the Sun and the smallest one in the Solar System—it's only a bit larger than our Moon. The planet's name has nothing to do with the liquid metal since it was named after the Roman messenger god, Mercury

VENUS

Venus has a beautiful name and is the second planet from the Sun. It's terribly hot—even hotter than Mercury—and its atmosphere is extremely poisonous. It's the second-brightest natural object in the night sky after the Moon

Maybe you need to divide the content

Mercury is the closest planet to the Sun and the smallest one in the Solar System—it's only a bit larger than our Moon. The planet's name has nothing to do with the liquid metal since it was named after the Roman messenger god, Mercury

Venus has a beautiful name and is the second planet from the Sun. It's terribly hot—even hotter than Mercury—and its atmosphere is extremely poisonous. It's the second-brightest natural object in the night sky after the Moon

You could use three columns, why not?

MARS

Despite being red, Mars is a cold place, not hot. It's full of iron oxide dust, which gives the planet its reddish cast

JUPITER

It's a gas giant and the biggest planet in our Solar System. Jupiter is the fourth-brightest object in the sky

VENUS

Venus has a beautiful name and is the second planet from the Sun. It's terribly hot—even hotter than Mercury

A picture always reinforces the concept

Images reveal large amounts of data quickly, so remember: use an image instead of long texts. Your audience will appreciate that.

Venus has a beautiful name and is the second planet from the Sun. It's terribly hot—even hotter than Mercury—and its atmosphere is extremely poisonous


A picture is worth a
thousand words

And here too!

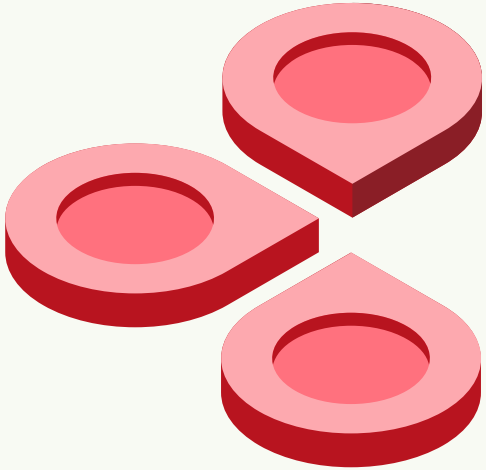
Images reveal large amounts of data quickly, so remember: use an image instead of long texts. Your audience will appreciate that

Awesome words

Because key words are great for catching your audience's attention

Infographics make your idea understandable...

JUPITER
It's a gas giant and the biggest planet in our Solar System



MARS
Despite being red, Mars is a cold place, not hot

VENUS
Venus has a beautiful name, but it's terribly hot

If you want to **modify this graph**, click on it, follow the link, change the data and replace it

MERCURY
Mercury is the closest planet to the Sun and is only a bit larger than our Moon

JUPITER
It’s a gas giant and the biggest planet in our Solar System

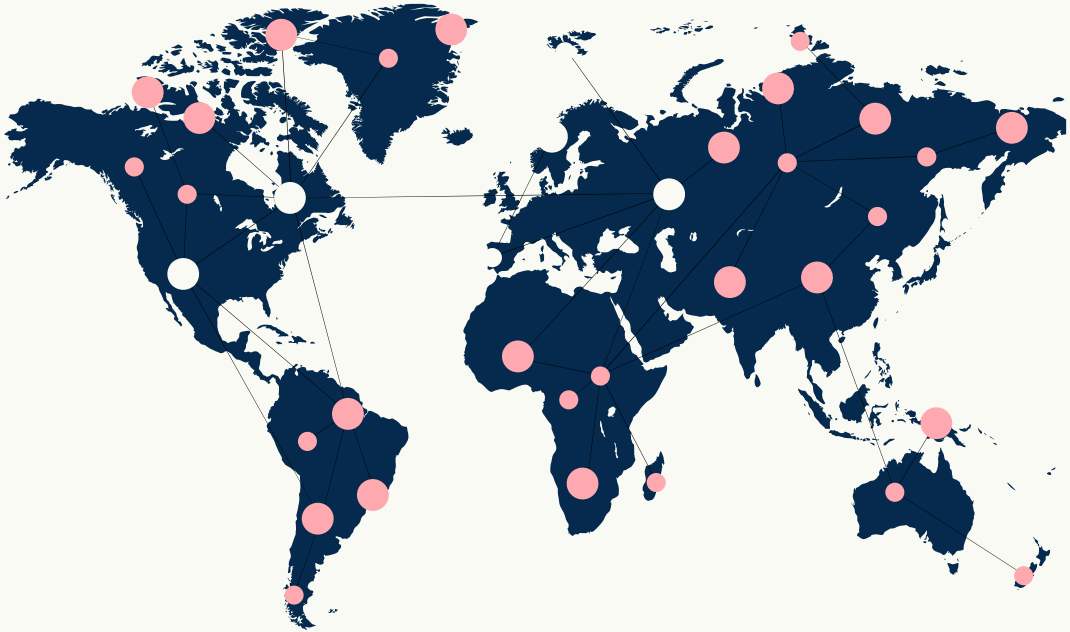
MARS
Despite being red, Mars is a cold place, not hot. The planet is full of iron oxide dust

VENUS
Venus has a beautiful name, but it’s terribly hot, even hotter than Mercury

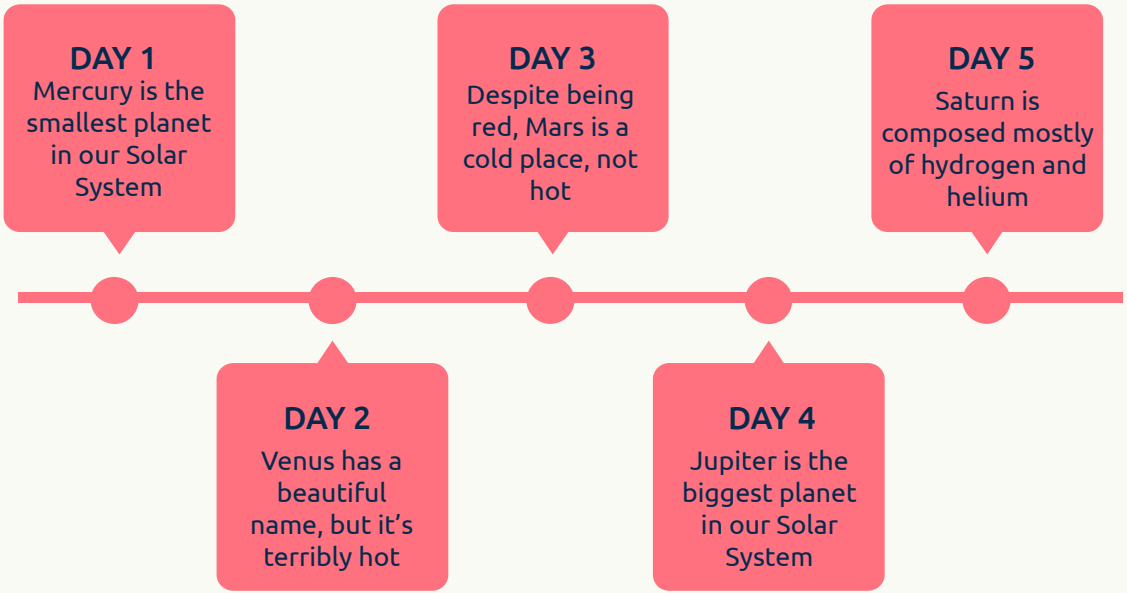
... and the same goes for tables

	MASS (earths)	DIAMETER (earths)	SURFACE GRAVITY (earths)
MERCURY	0,06	0,38	0,38
MARS	0,11	0,53	0,38
SATURN	95,2	9,4	1,16

This is a map



A timeline always works fine



4,498,300,000

Big numbers catch your audience's attention

333,000.00

earths is the Sun's mass

24h 37m 23s

is Jupiter's rotation period

386,000 km

is the distance between the Earth and the Moon

Sometimes, reviewing concepts is a good idea

MERCURY

Mercury is the closest planet to the Sun and is only a bit larger than our Moon

VENUS

Venus has a beautiful name, but it's terribly hot, even hotter than Mercury

MARS

Despite being red, Mars is a cold place, not hot. The planet is full of iron oxide dust

JUPITER

It's a gas giant and the fourth-brightest object in the sky

SATURN

Yes, this is the ringed one. It's a gas giant, composed mostly of hydrogen and helium

NEPTUNE

Neptune is the farthest planet in our Solar System and the fourth-largest

Sometimes, reviewing concepts is a good idea

VENUS

Venus has a beautiful name and is the second planet from the Sun. It's terribly hot

MARS

Despite being red, Mars is a cold place, not hot. The planet is full of iron oxide dust

SATURN

Yes, this is the ringed one. It's a gas giant, composed mostly of hydrogen and helium

NEPTUNE

Neptune is the farthest planet in our Solar System and also the fourth-largest

Your services

MERCURY

Mercury is the closest planet to the Sun and the smallest one in the Solar System—it's only a bit larger than our Moon. The planet's name has nothing to do with the liquid metal since it was named after the Roman messenger god, Mercury

VENUS

Venus has a beautiful name and is the second planet from the Sun. It's terribly hot—even hotter than Mercury—and its atmosphere is extremely poisonous. It's the second-brightest natural object in the night sky after the Moon

More services

MARS

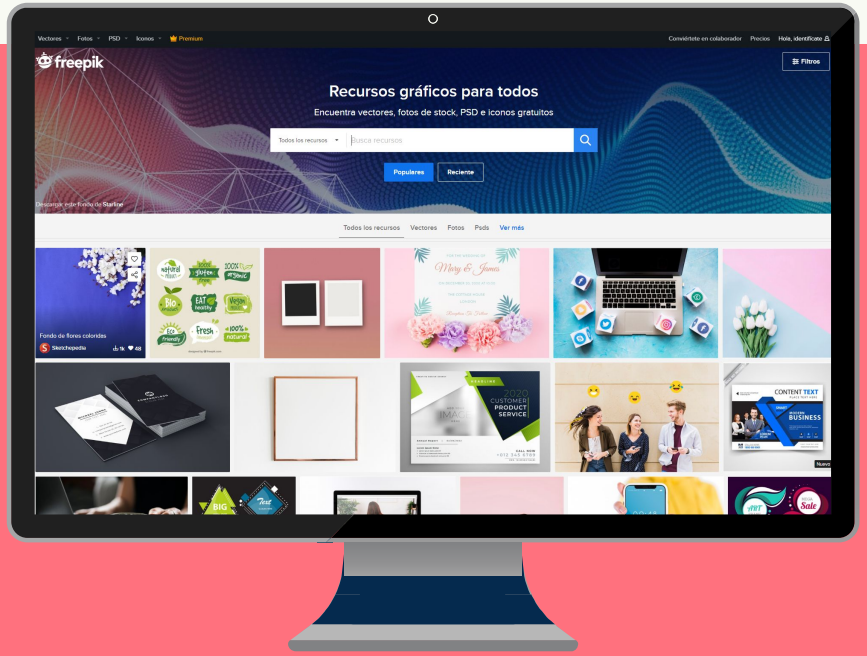
Despite being red, Mars is a cold place, not hot. It's full of iron oxide dust, which gives the planet its reddish cast

JUPITER

It's a gas giant and the biggest planet in our Solar System. Jupiter is the fourth-brightest object in the sky

VENUS

Venus has a beautiful name and is the second planet from the Sun. It's terribly hot—even hotter than Mercury

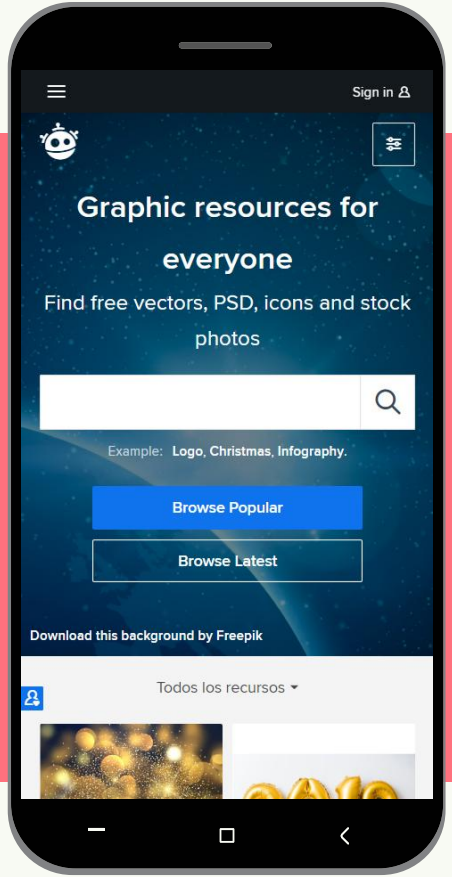


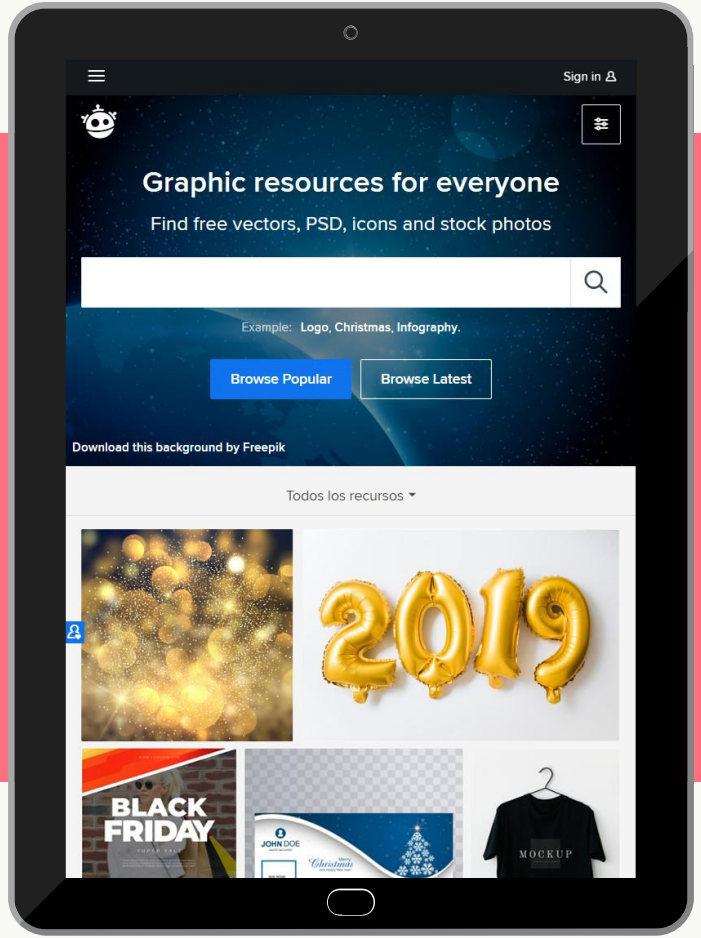
Desktop Software

You can replace the image on the screen with your own work. Just delete this one, add yours and send it to the back

Mobile Web

You can replace the image on the screen with your own work. Just delete this one, add yours and send it to the back





Tablet App

You can replace the image on the screen with your own work. Just delete this one, add yours and send it to the back

Talk about your TEAM

GRAPHIC DESIGNER

Some words about this person

CREATIVE

Some words about this person

PHOTOGRAPHER

Some words about this person

VIDEO EDITOR

Some words about this person

Thanks!

Does anyone have any questions?

addyouremail@freepik.com
+91 620 421 838
yourcompany.com

Credits

This is where you give credit to the ones who are part of this project.

Did you like the resources on this template? Get them for **free** at our other websites.

Presentation template by Slidesgo

Icons by Flaticon

Images & infographics by Freepik

Author introduction slide photo created by **katemangostar** - Freepik.com

Big image slide photo created by **jcomp** - Freepik.com

Text & Image slide photo created by **rawpixel.com** - Freepik.com

Text & Image slide photo created by **Freepik**

Instructions for use

In order to use this template, you must credit **Slidesgo** by keeping the Credits slide.

You are allowed to:

- Modify this template.
- Use it for both personal and commercial projects.

You are not allowed to:

- Sublicense, sell or rent any of Slidesgo Content (or a modified version of Slidesgo Content).
- Distribute Slidesgo Content unless it has been expressly authorized by Slidesgo.
- Include Slidesgo Content in an online or offline database or file.
- Offer Slidesgo templates (or modified versions of Slidesgo templates) for download.
- Acquire the copyright of Slidesgo Content.

For more information about editing slides, please read our FAQs:
https://slidesgo.com/faqs

Fonts & colors used

This presentation has been made using the following fonts:

Ubuntu
(https://fonts.google.com/specimen/Ubuntu)

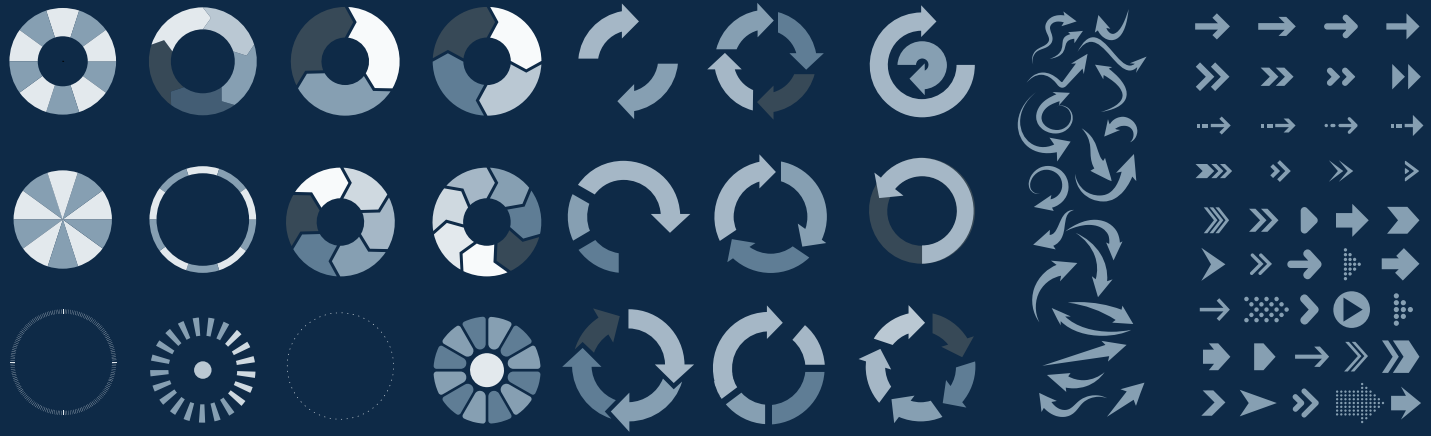
#ff737d

#fda9af

#ffffffff

Use our editable graphic resources...

You can easily resize these resources keeping the quality. To change the color just ungroup the resource and click on the object you want to change. Then click on the paint bucket and select the color you want. Don’t forget to group the resource again when you’re done.



...and our set of editable icons

You can resize these icons keeping the quality.

You can change the stroke and fill color; just select the icon and click on the paint bucket/pen.

Business Icons

Avatar Icons

Creative Process Icons

Educational Process Icons

Help & Support Icons

Medical Icons

Nature Icons

Performing Arts Icons

SEO & Marketing Icons

Teamwork Icons


slidesgo