



Embedding Domain Knowledge in DNN

AI Shot #2 @ VISUM

Kelwin Fernandes, Ph.D.
CEO at nilg.ai
kelwin@nilg.ai



01

The Learning Spectrum

AI shot

02

Embedding Domain Knowledge

AI shot

03

Semi-Supervised Learning

AI shot

04

AI use case

01

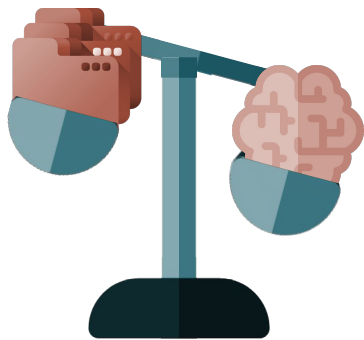
The good
old days



The good old days

Before and After Deep Learning

Before Deep Learning

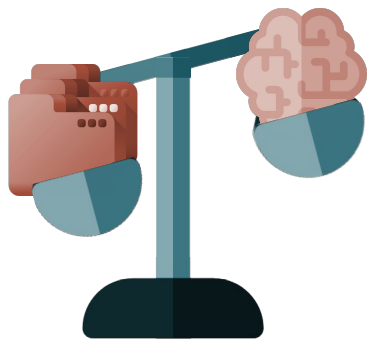


- Tons of focus on **feature engineering** and **domain understanding**.
- Manual-intensive task.
- Do you need a COVID-19 detection model?
 - Sure! give me 10 Ph.D., a budget for 5 years, 1K images.
 - Then, I'll come close to a PoC that can only be used under these strictly designed settings.

The good old days

Before and After Deep Learning

After Deep Learning



- The new popular kid arrived to the school.
- Getting models for new applications “for-free”.
- Do you need a COVID-19 detection model?
 - Sure! give me 1 month, a monkey capable of pressing two buttons (fit+predict) and 1M data points.
 - Then, I'll have a model that seems to be working better than humans... at least on silico.

The good old days

Before and After Deep Learning

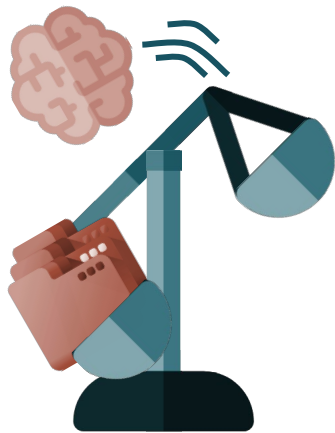


After Deep Learning

- The new popular kid arrived to the school.
- Getting models for new applications “for-free”.
- Do you need a COVID-19 detection model?
 - Sure! give me 1 month, a monkey capable of pressing two buttons (fit+predict) and 1M data points.
 - Then, I'll have a model that seems to be working better than humans... at least on silico.
- So, people got mad...

The good old days

Before and After Deep Learning



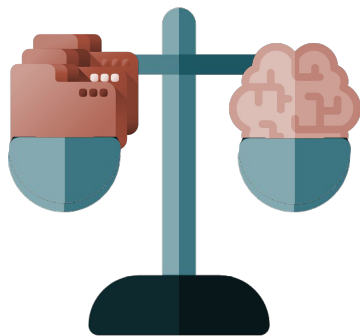
After Deep Learning

- The new popular kid arrived to the school.
- Getting models for new applications “for-free”.
- Do you need a COVID-19 detection model?
 - Sure! give me 1 month, a monkey capable of pressing two buttons (fit+predict) and 1M data points.
 - Then, I'll have a model that seems to be working better than humans... at least on silico.
- So, people got mad... up to the point where we almost discard human experience as a valid source of knowledge.

The good old days

Before and After Deep Learning

Embedding Domain Knowledge in DNN



- Not so bald nor with two wigs
 - (a bite of venezuelan wisdom)
- We can have a mid-point where:
 - We understand the business.
 - We understand the tech.
 - We know how to elevate the tech to the business.
 - Instead of dumbing down the business to fit the tech.

02

Embedding Domain Knowledge in DNN



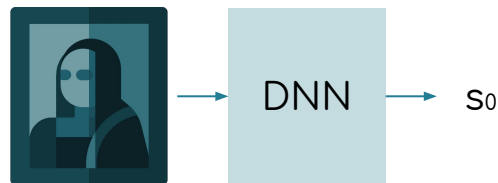
We will discuss multiple
examples and how to build
tailored DNN for each case...

Fasten your seatbelts.
Please try this at home!

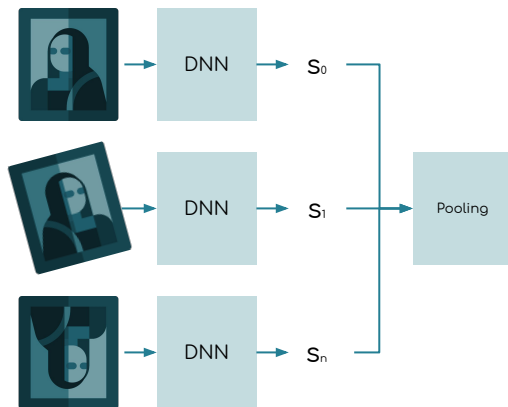


Adding Invariances: The easy way

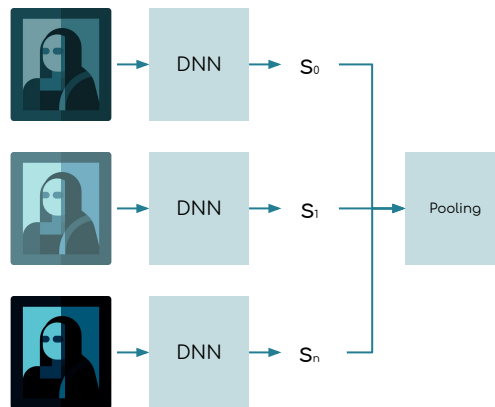
Synthetic (or natural) Data
Augmentation at Inference



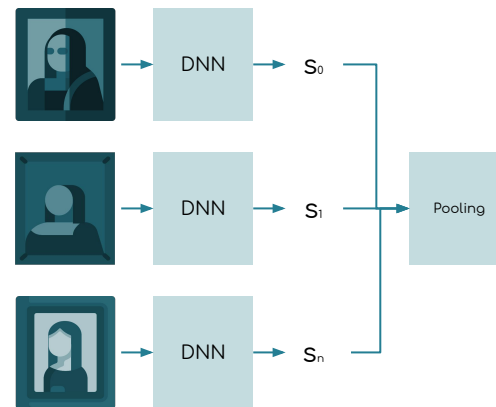
Flips/Rotations



Illumination/Contrast



Multiple Views/Poses



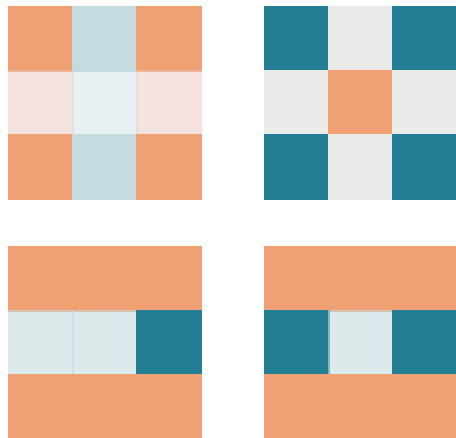
Adding Invariances: The hard way

Local Preprocessing or...

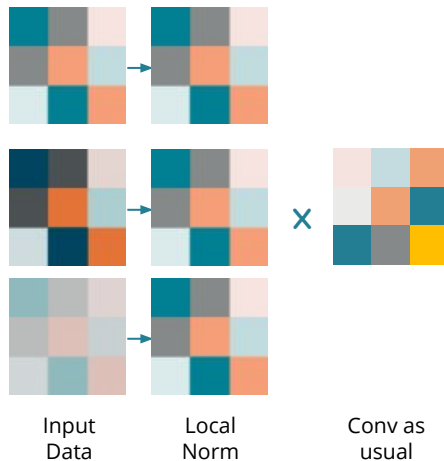
Building tailored convolutional
kernels/architectures



Flips/Rotations



Illumination/Contrast



Multiple Views/Poses

Ask Geoffrey Hinton
about
capsule neural
networks

Adding Invariances

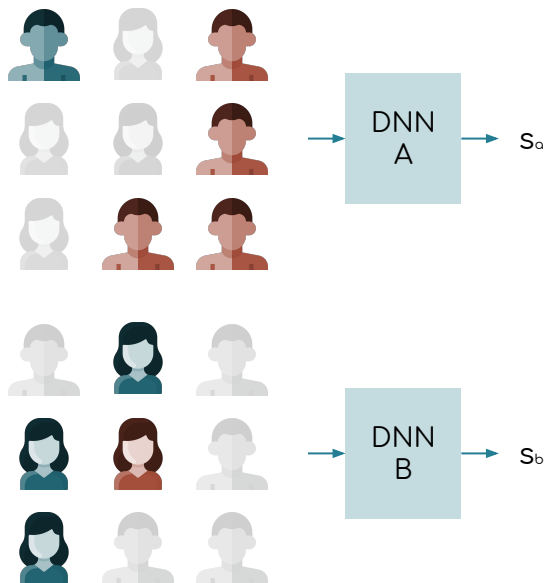
Group invariance:

race, gender, country, deep-learning-framework-preference

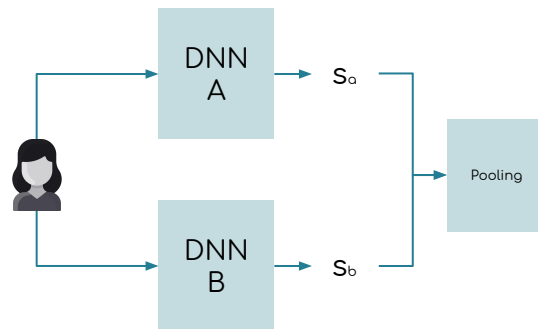
Biased Dataset



Train a model per group

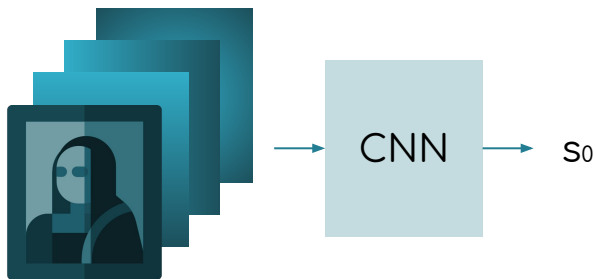
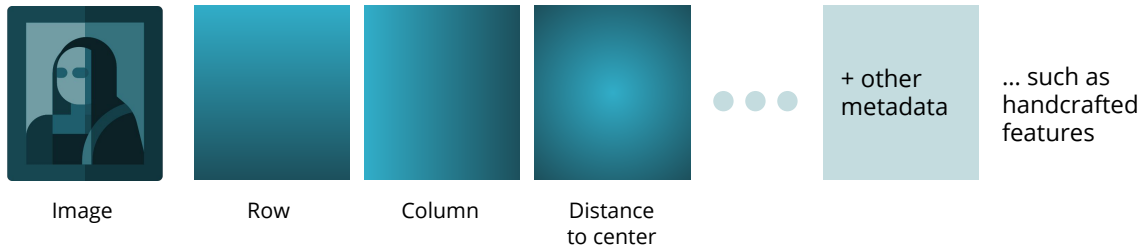


Avg. voting as prediction



Removing Invariances

- Remove translational invariance from convolutional kernels
 - Add pixel location as additional channels

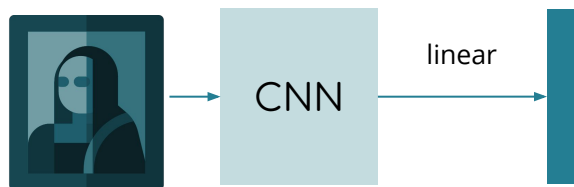


Note: you can have two sets of kernels, one with and one without spatial coordinates so you get the best from the two worlds.

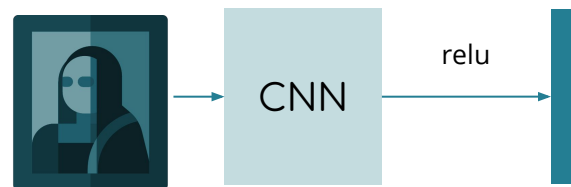
Social Media Reactions

Predict: likes, love, care, haha, wow, sad, angry

First attempt:
naive regression



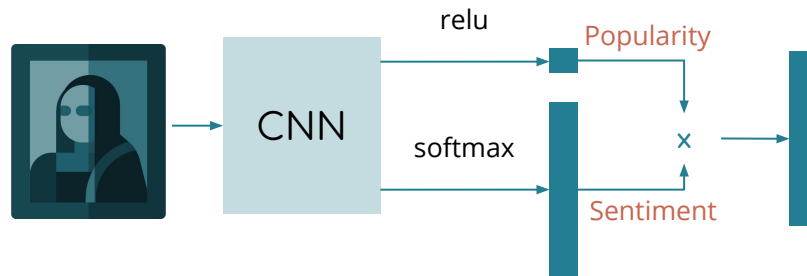
Second attempt:
regression knowing you cannot downvote



Third attempt:

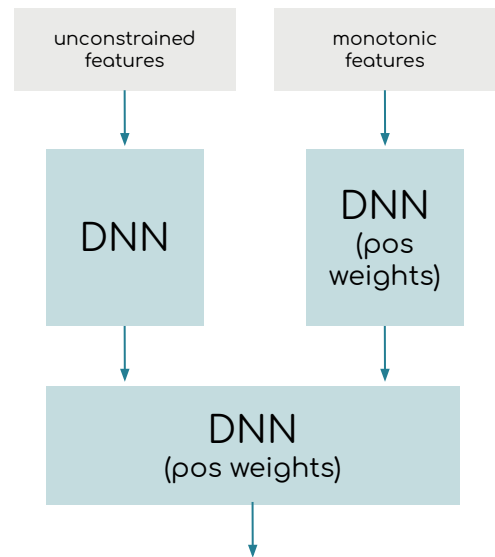
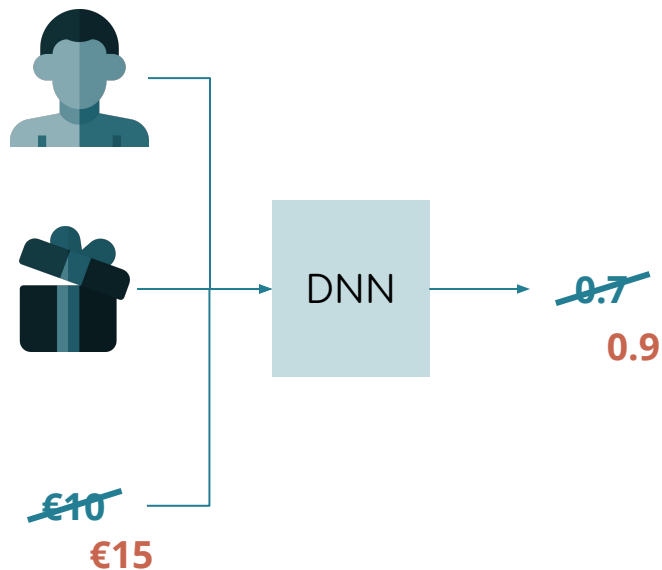
Ease the task by
learning 2 simpler tasks:

- Popularity
- Sentiment



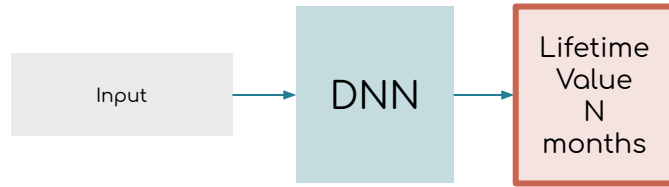
Predict if a customer will buy a product

Forcing monotonic Behavior



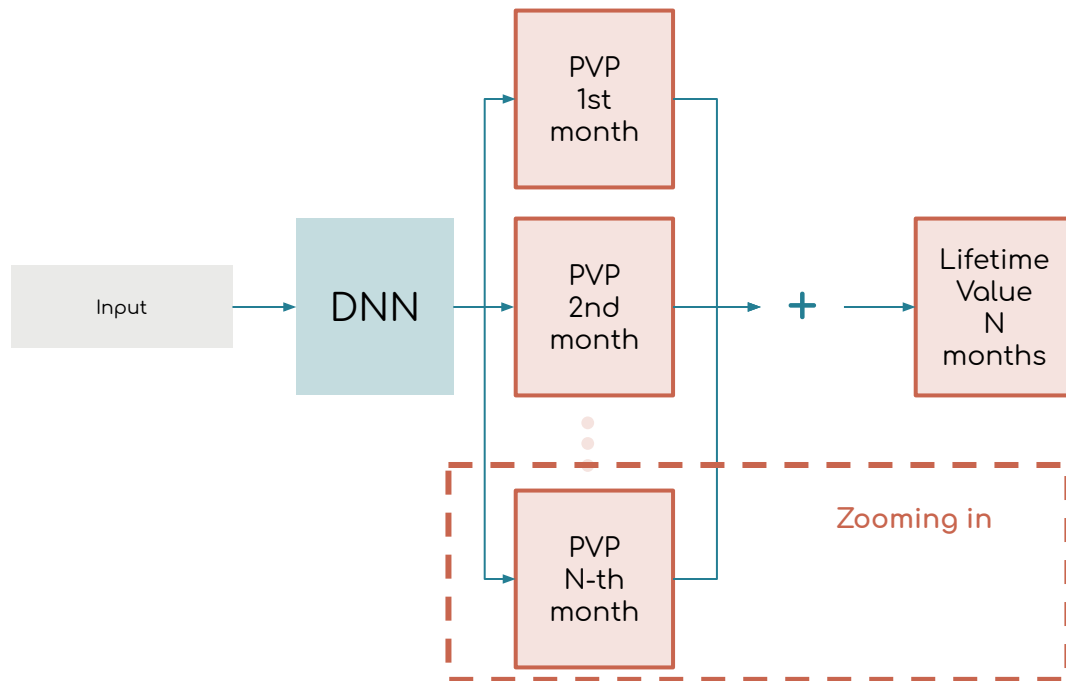
Customer Lifetime Value

When harder is simpler



Customer Lifetime Value

When harder is simpler



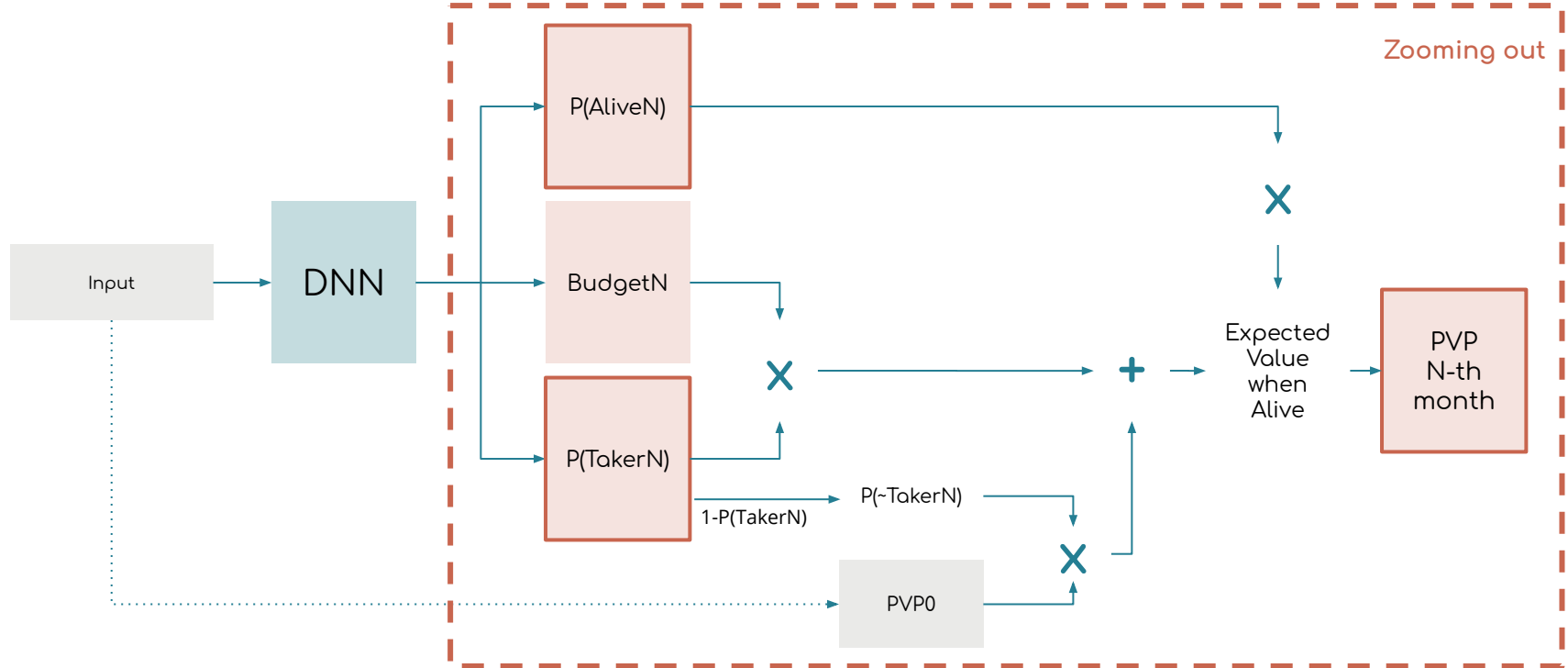
$$\text{Expected PVP} = \text{Prob}(\text{Alive}) * \text{TotalBudget} + (1 - \text{Prob}(\text{Alive})) * 0$$

$$\text{TotalBudget} = \text{Prob}(\text{Taker}) * \text{Budget} + (1 - \text{Prob}(\text{Taker})) * \text{InitialPVP}$$

$$\text{Expected PVP} = \text{Prob}(\text{Alive}) * (\text{Prob}(\text{Taker}) * \text{Budget} + (1 - \text{Prob}(\text{Taker})) * \text{InitialPVP})$$

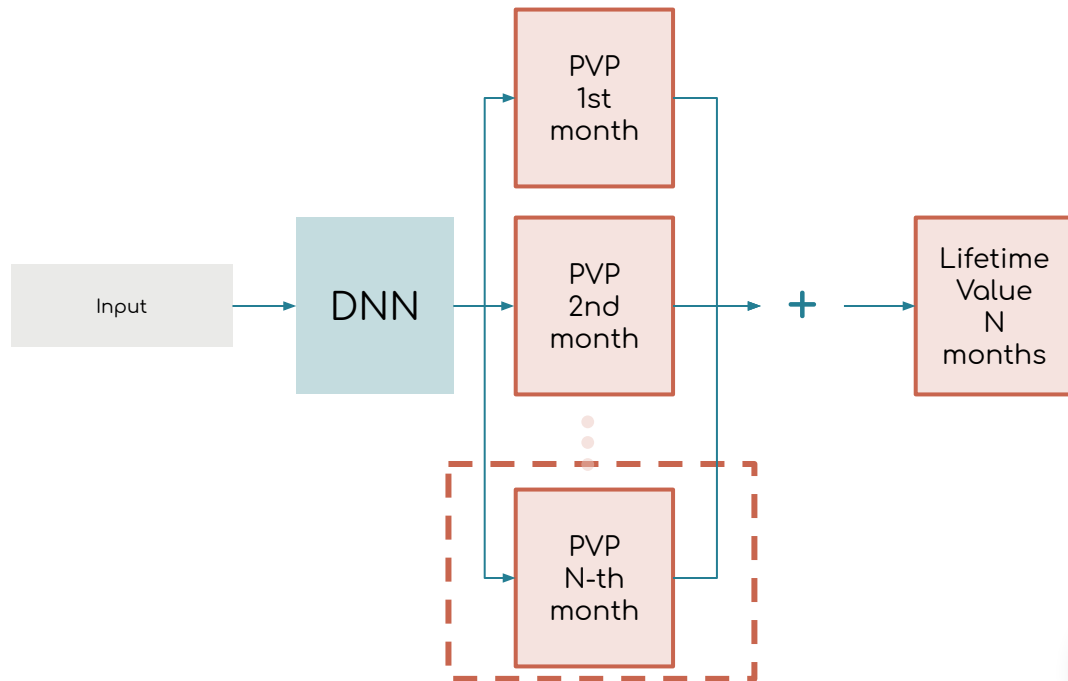
Customer Lifetime Value

When harder is simpler



Customer Lifetime Value

When harder is simpler



Why adding complexity and investing that much effort?

- Additional tasks act as regularizers.
- Giving additional supervision of the internal business processes:
 - Facilitates the discovery of features relevant for those tasks.
 - Avoids catastrophic failures due to unmatching business-machine rules.
- We reduced the MAE by 50% ;-)



<https://nilg.ai/blog/202004/embedding-domain-knowledge-for-estimating-customer-lifetime-value/>

Conclusion



- Deep Neural Networks are not the black-box monster they told you.
- They can actually be a good dog if you know how to train (or breed) them!
- Domain knowledge isn't (always) bad.
 - It can help you to learn with **less data**.
 - It can help you avoid **catastrophic failure**.
 - So you can rest tonight! Assured that your Project Manager won't call you at midnight because your model decided to sell cars at \$0 to maximize hit rate.

Thank you!
Questions?





Embedding Domain Knowledge in DNN

AI Shot #2 @ VISUM

Kelwin Fernandes, Ph.D.
CEO at nilg.ai
kelwin@nilg.ai

