Math-Industry Day
IHP - Paris
November 23th, 2016

Towards new Machine Learning tools based on Ode and Pde models

Emmanuel Frénod

Laboratoire de Mathématiques de Bretagne Atlantique (UMR6205) - Université Bretagne-Sud - Vannes
&
See-d

DATA

Data

Phase	1			2			3		
Period	1 → 16			17 → 22			23 → 43		
Lipid proportion	20%			17%			18%		
Protein proportion	40%			45%			48%		
Consumption	2785	…	2834	1983	…	2013	1675	…	1669
Weight gain	2078	…	2107	1430	…	1468	987	…	965

3

Data

Phase	1	2	3
Period	1 → 16	17 → 22	23 → 43
Lipid proportion	20%	17%	18%
Protein			
Consum			
Weight			

Phase	1	2	3
Period	1 → 16	17 → 22	23 → 43
Lipid proportion	23%	18%	20%
Protein			
Consum			
Weight			

Phase	1			2			3		
Period	1 → 16			17 → 22			23 → 43		
Lipid proportion	22%			16%			19%		
Protein proportion	36%			47%			46%		
Consumption	2786	…	2827	1890	…	1976	1665	…	1678
Weight gain	2059	…	2113	1434	…	1468	978	…	965

…

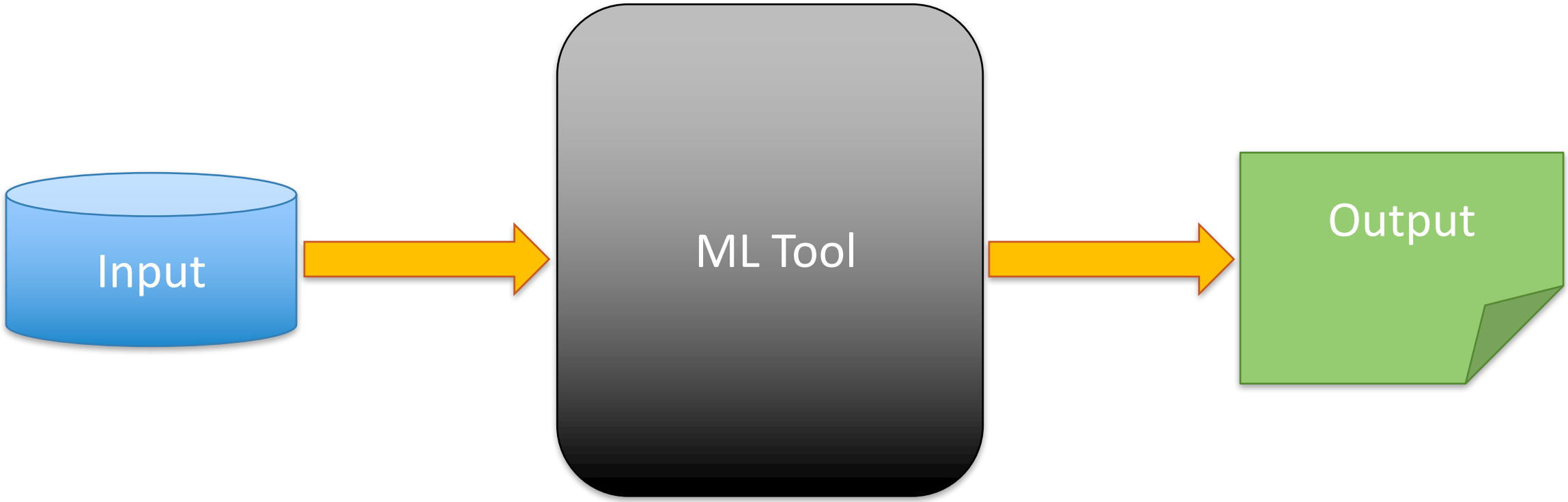
Machine Learning

Objective : Learn to forecast

5

Machine Learning

Objective : Learn to forecast



6

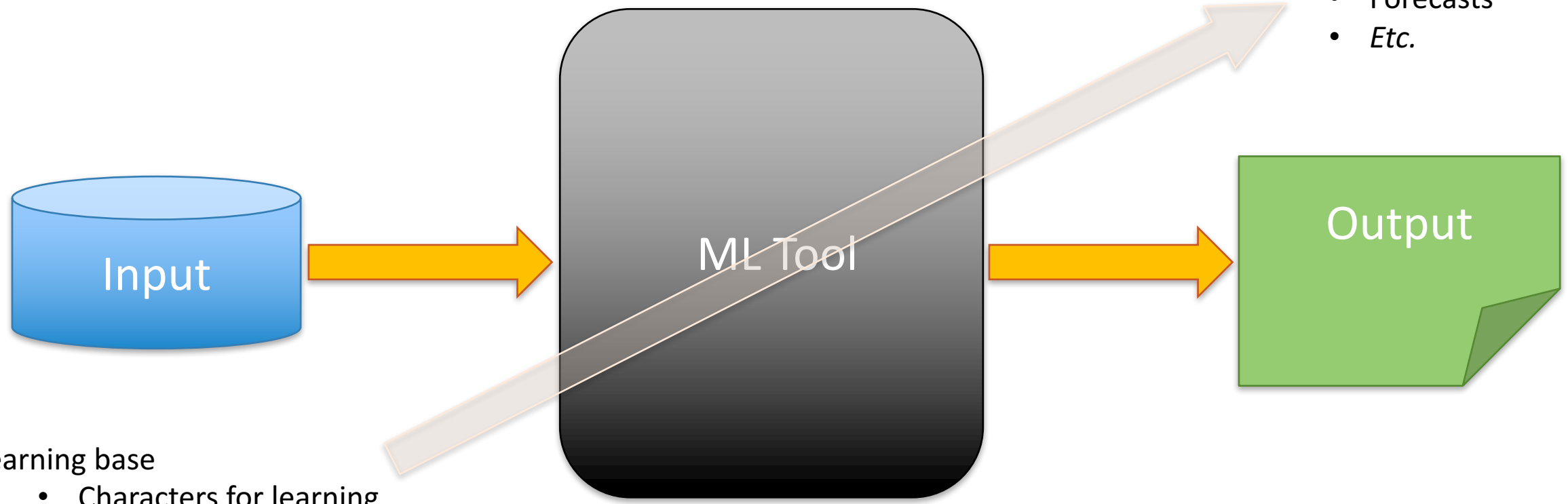
Machine Learning

For our problem

7

Machine Learning

Generic working



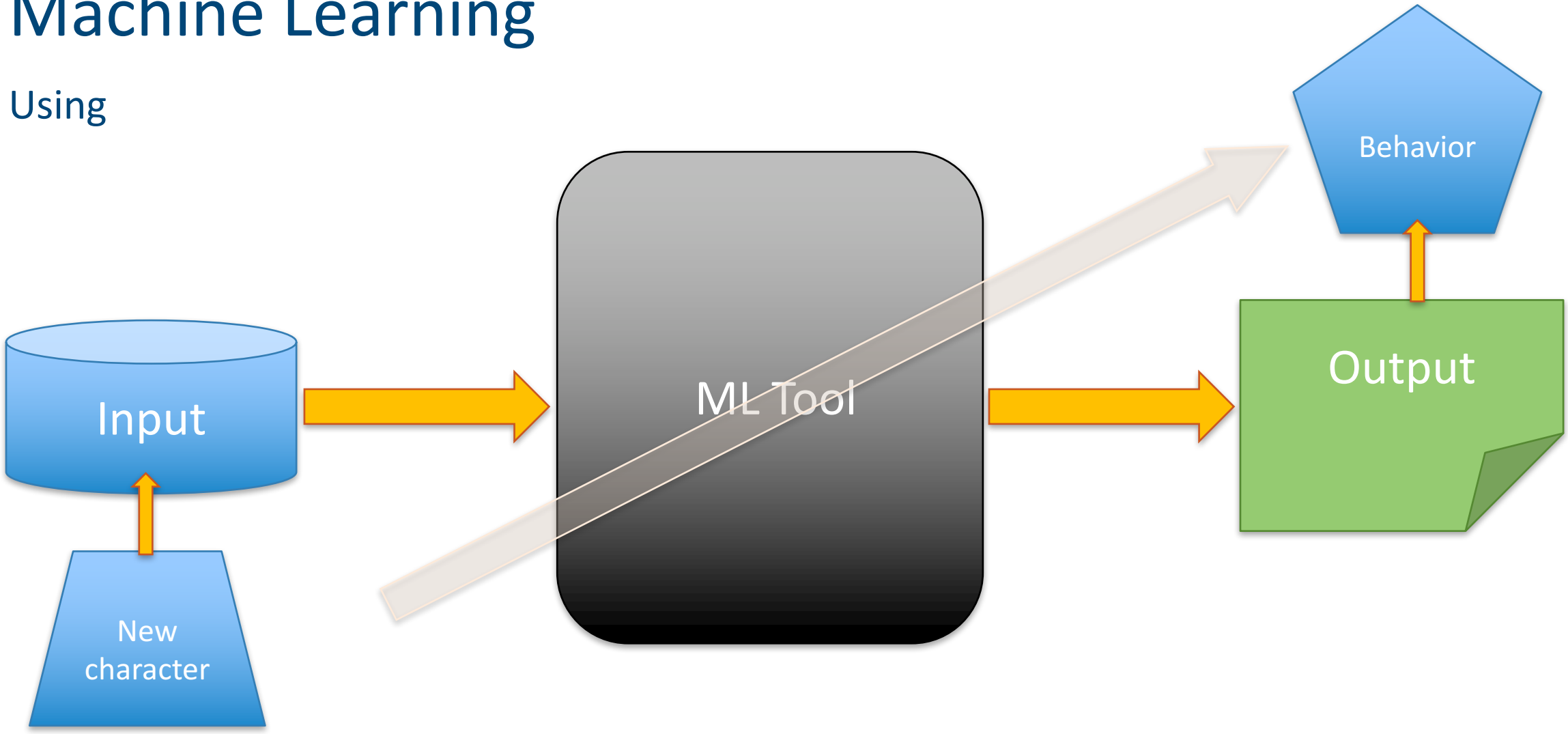
- Groups
- Correlations
- Forecasts
- *Etc.*

Learning base

- Characters for learning
- Characters for tests

Machine Learning

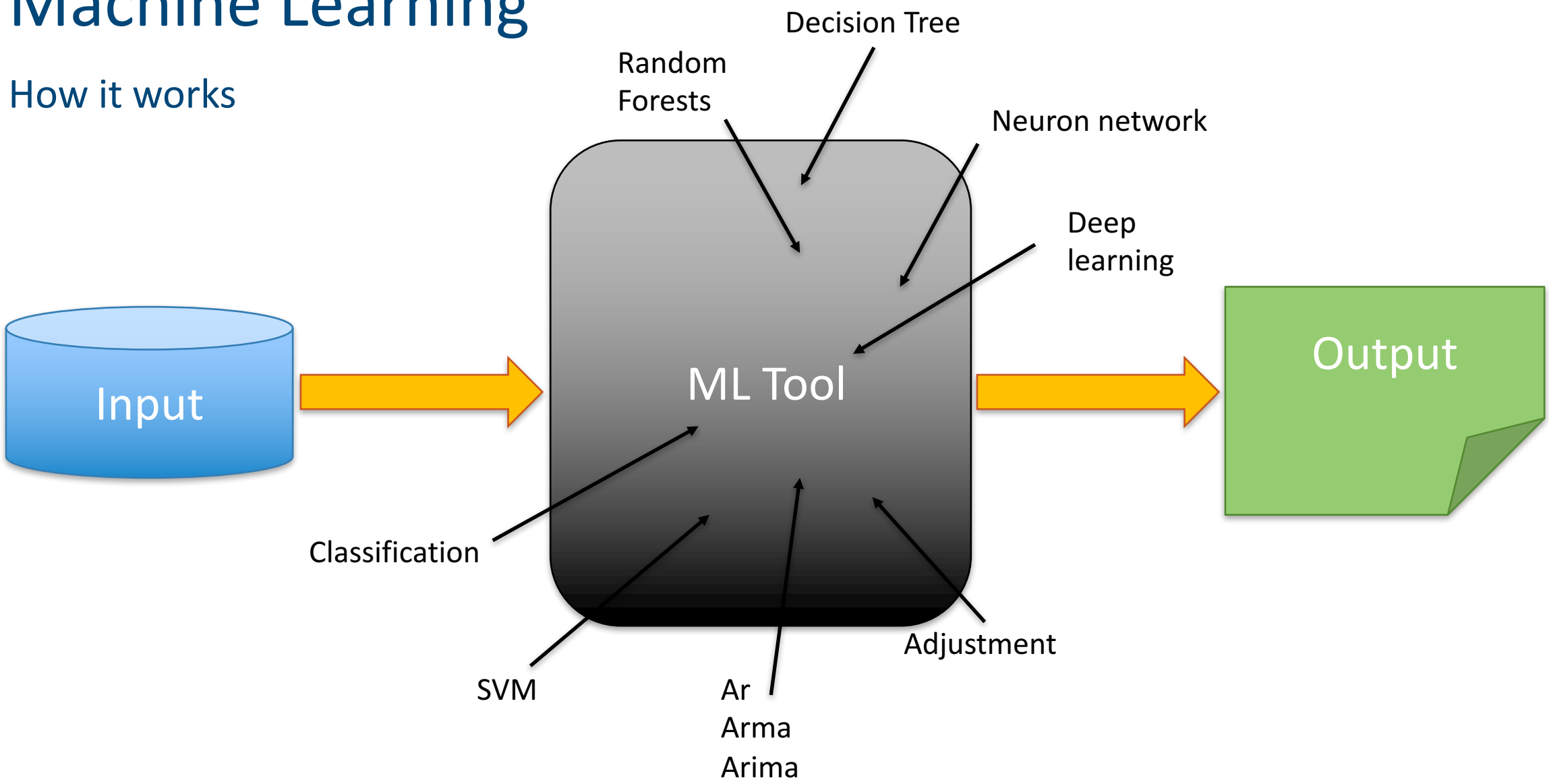
Using



9

Machine Learning

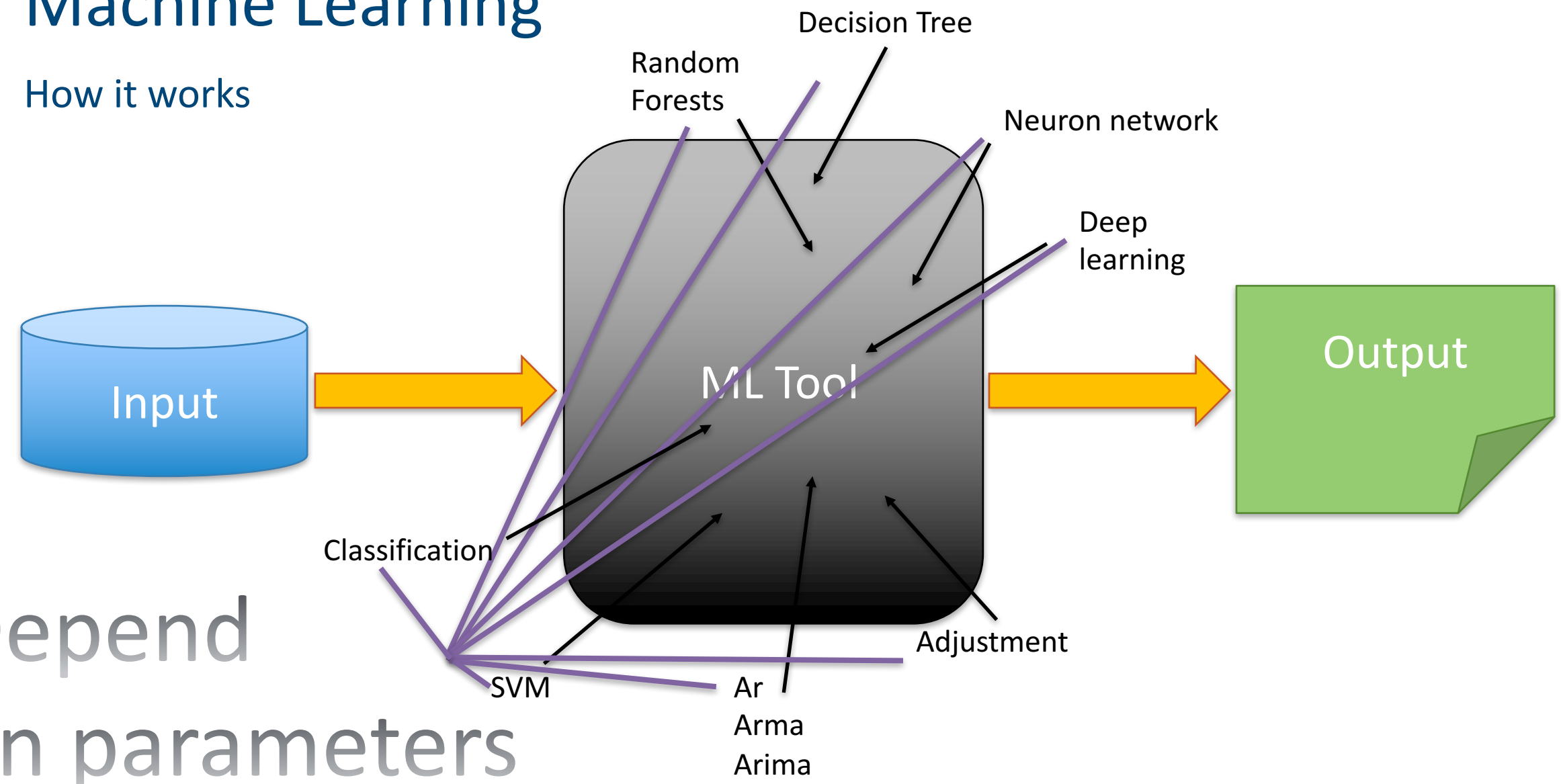
How it works



10

Machine Learning

How it works



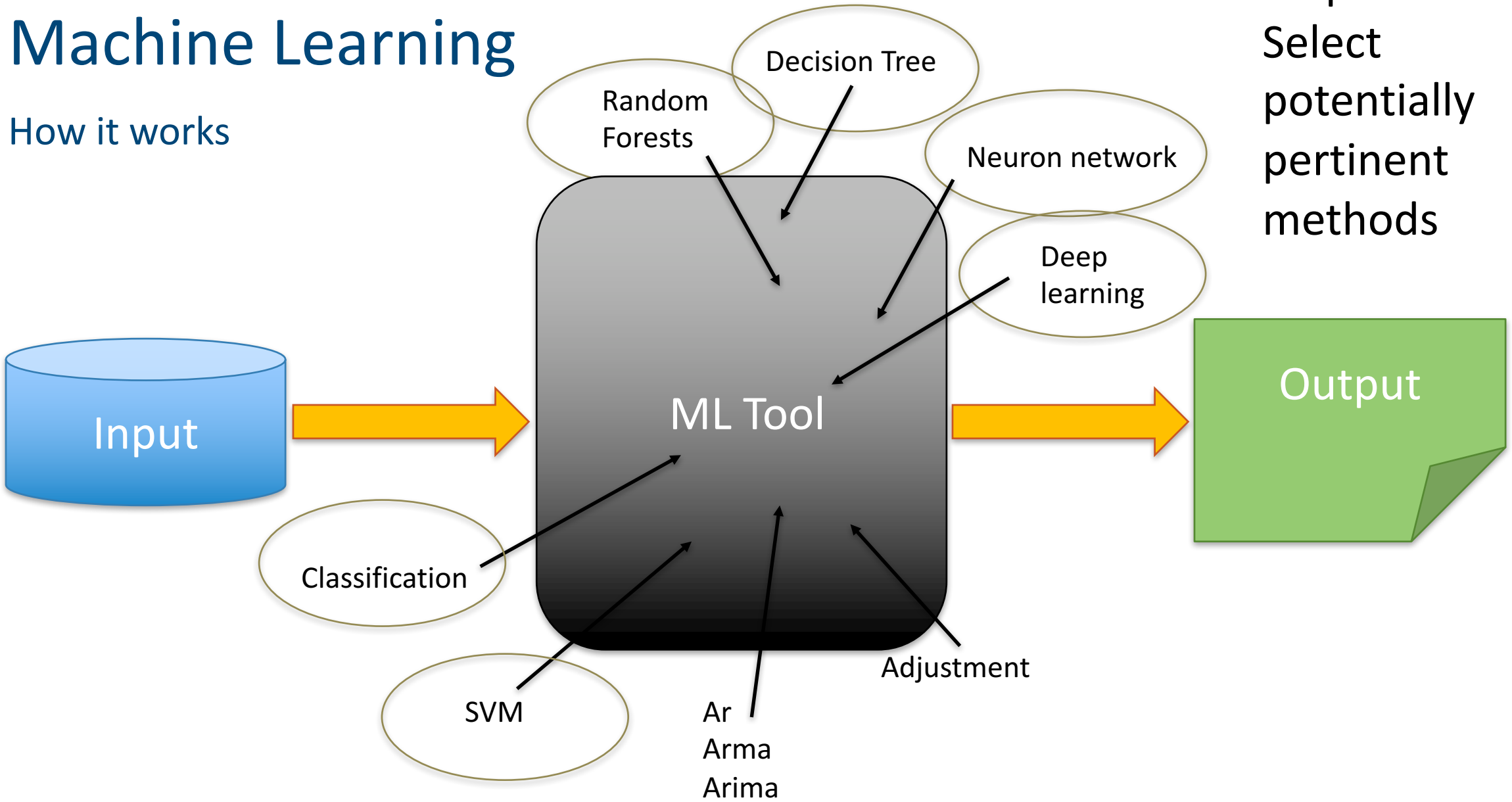
Depend on parameters

11

Machine Learning

How it works

Step 1 :
Select potentially pertinent methods



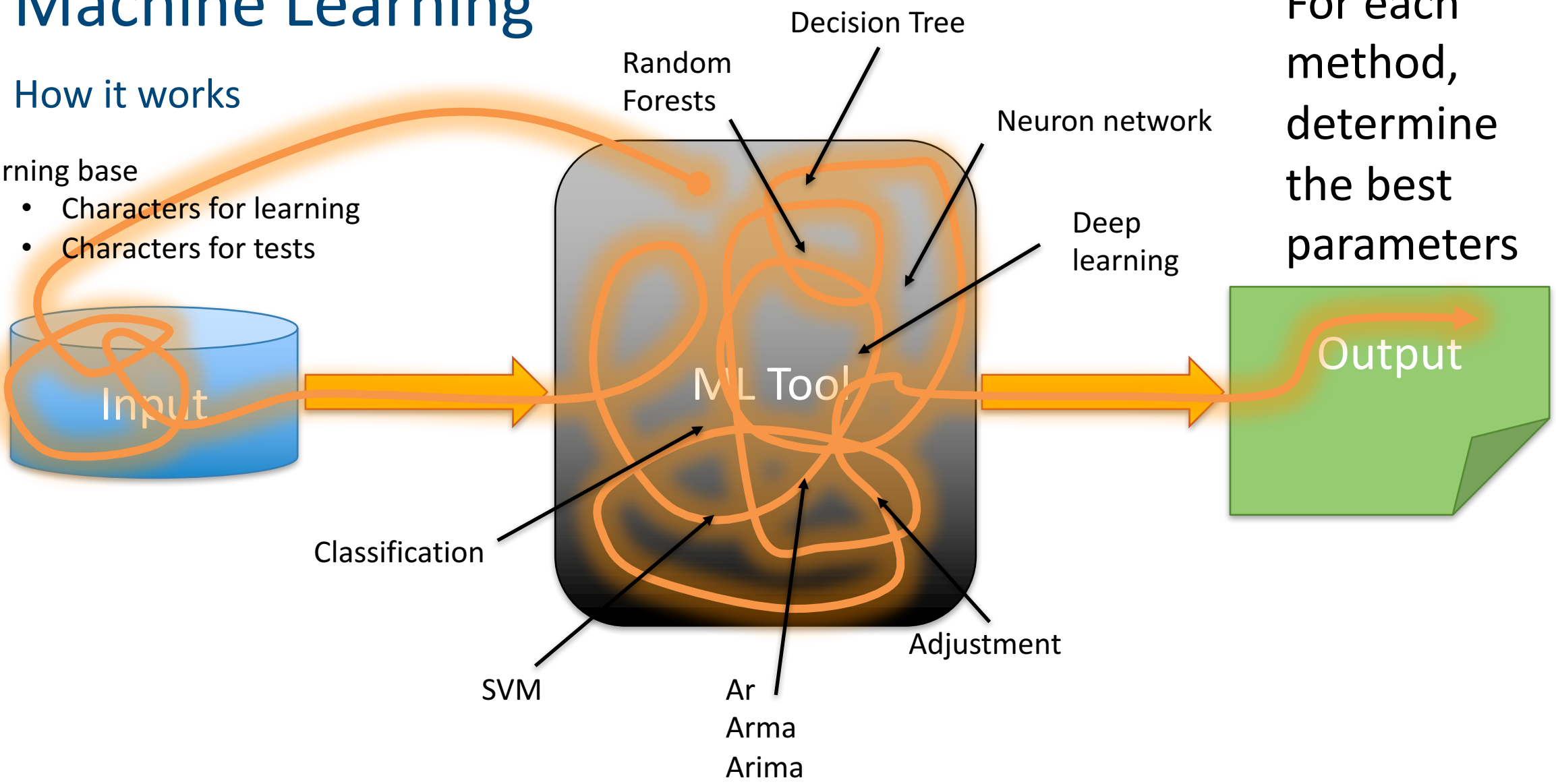
12

Machine Learning

How it works

Learning base

- Characters for learning
- Characters for tests



Step 2 :
For each method, determine the best parameters

13

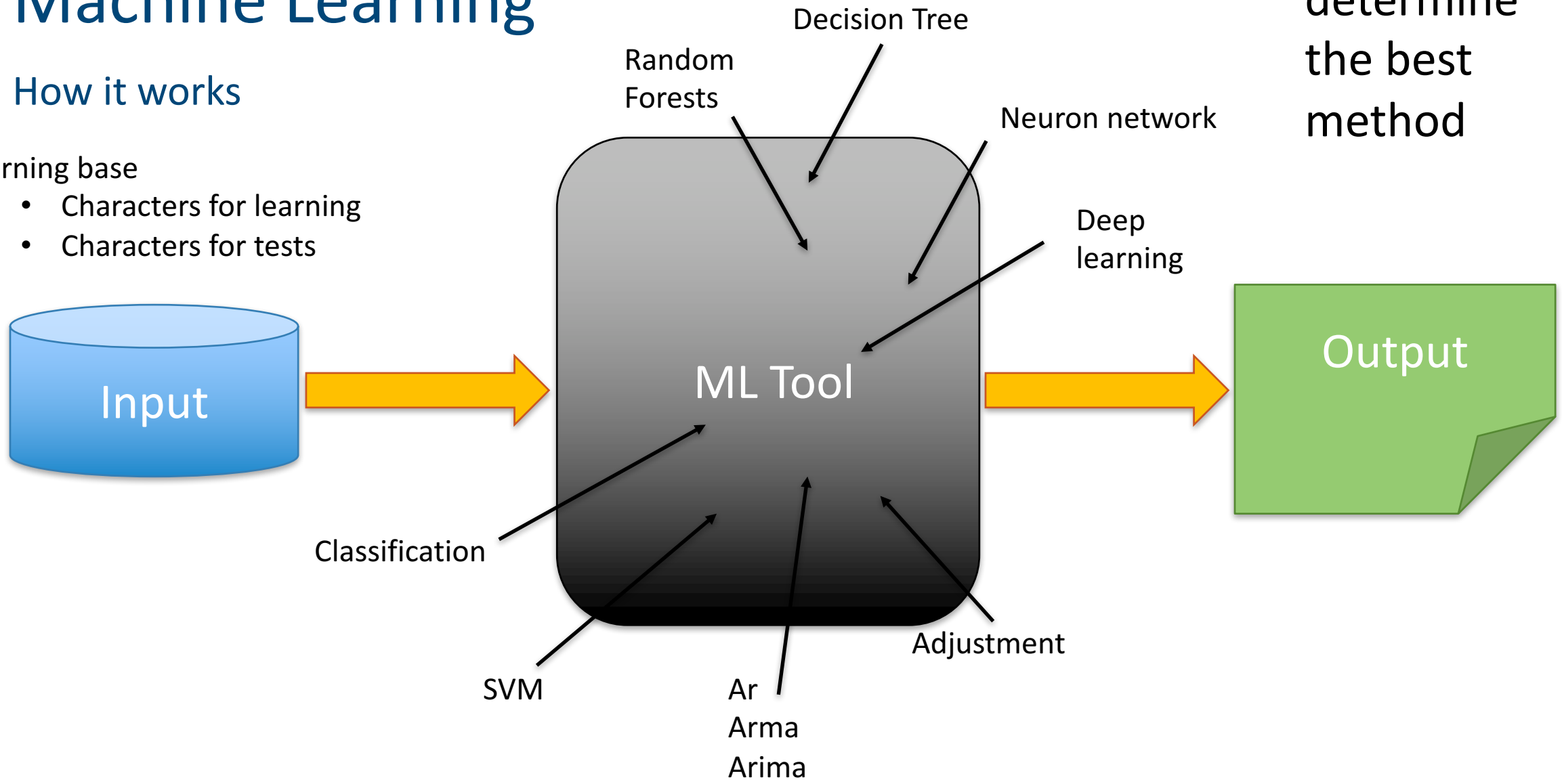
Machine Learning

Step 3 : determine the best method

How it works

Learning base

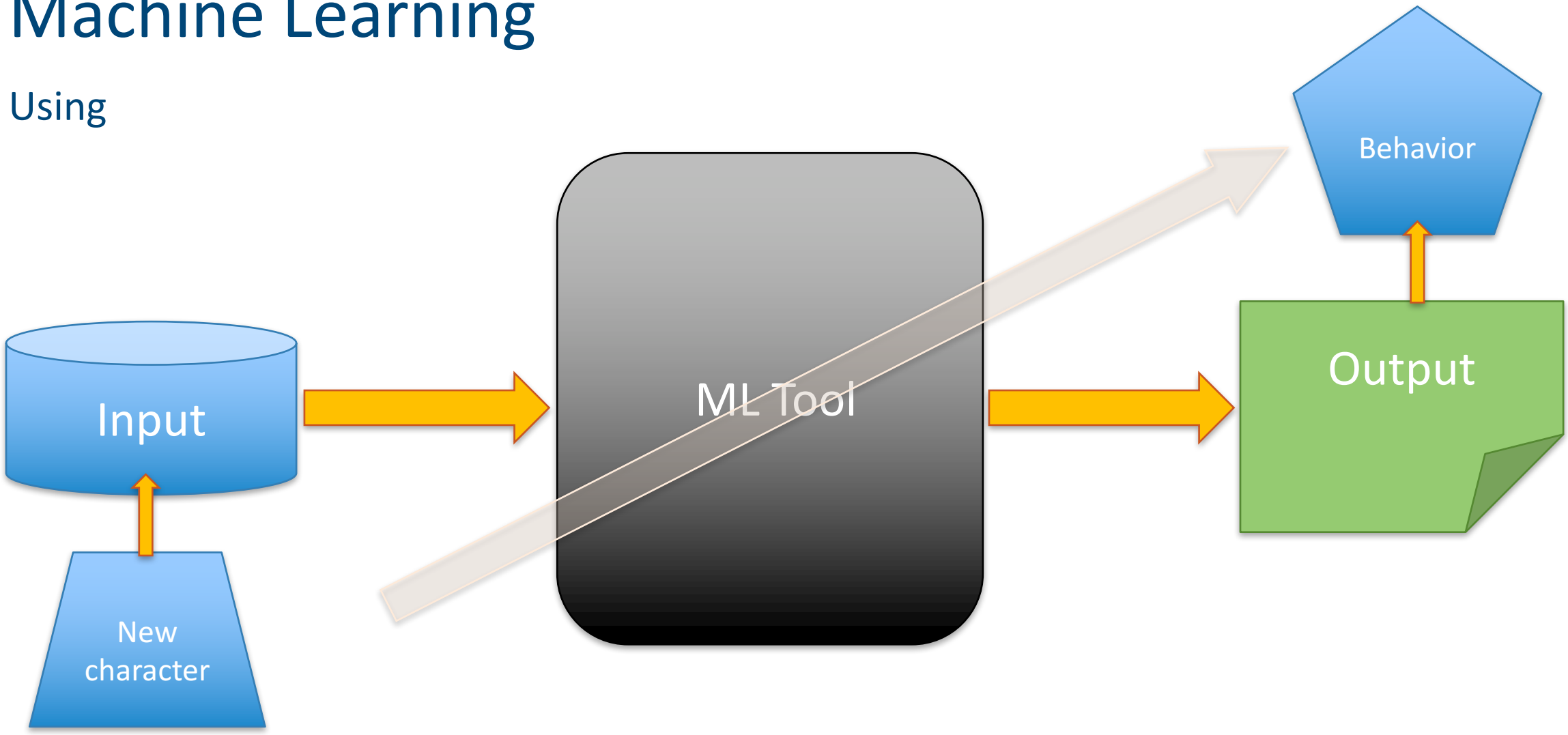
- Characters for learning
- Characters for tests



14

Machine Learning

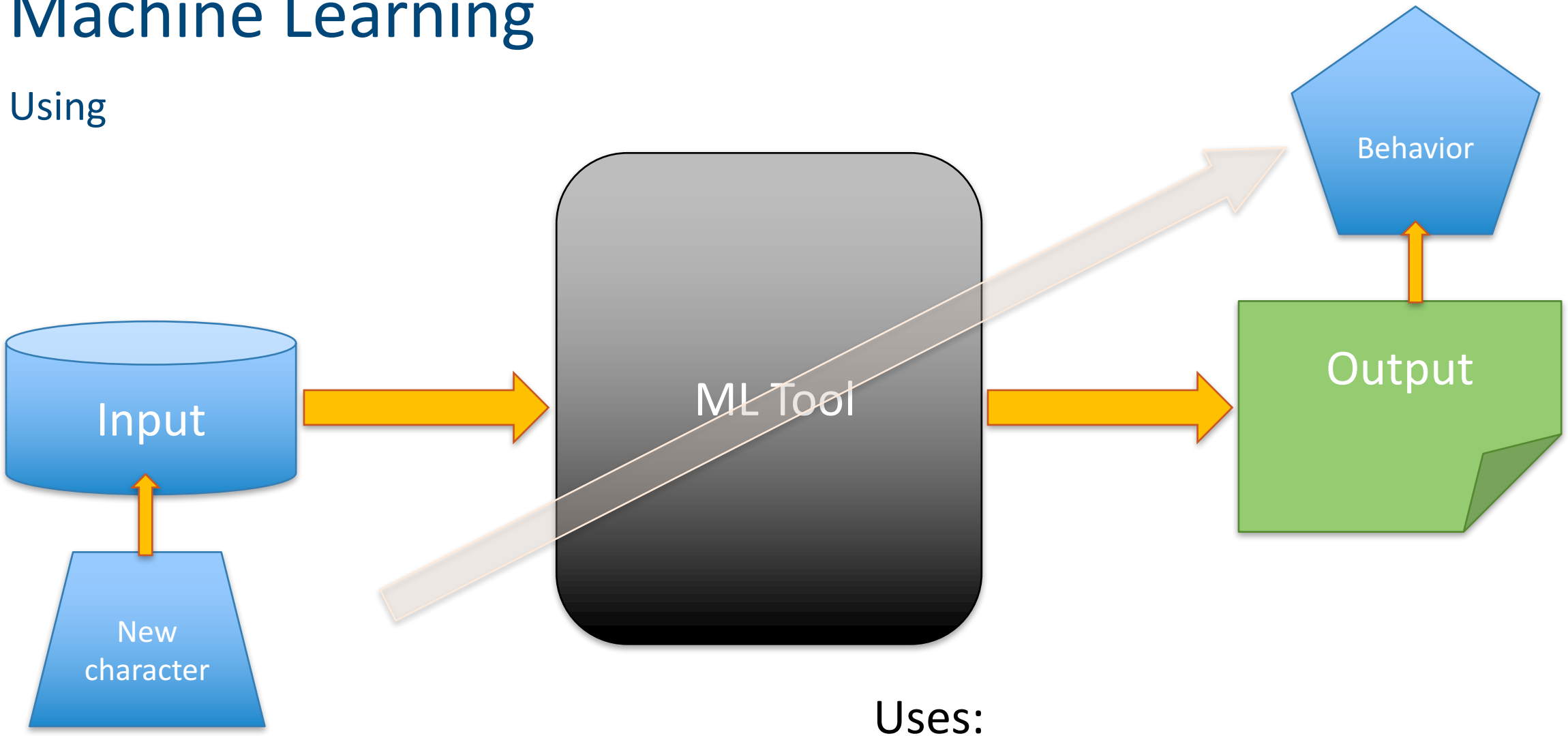
Using



15

Machine Learning

Using



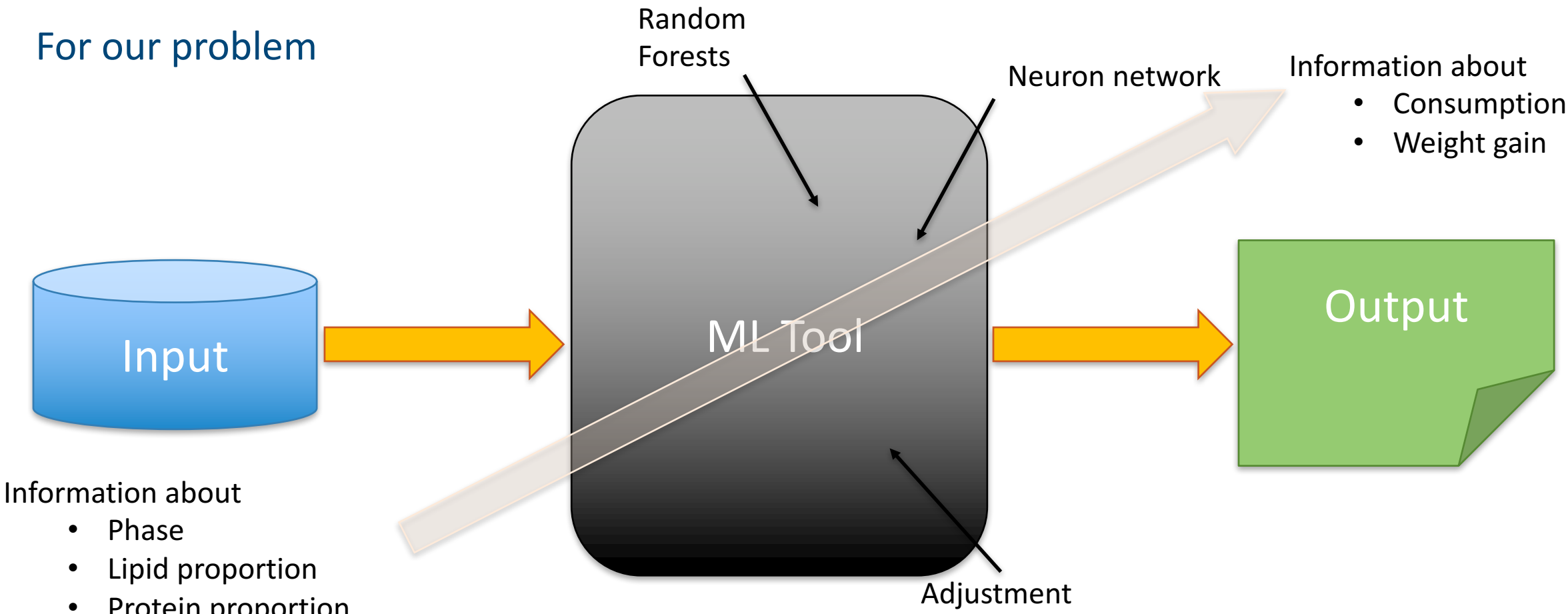
Uses:

- The best method
- With the right parameters

16

Machine Learning

For our problem

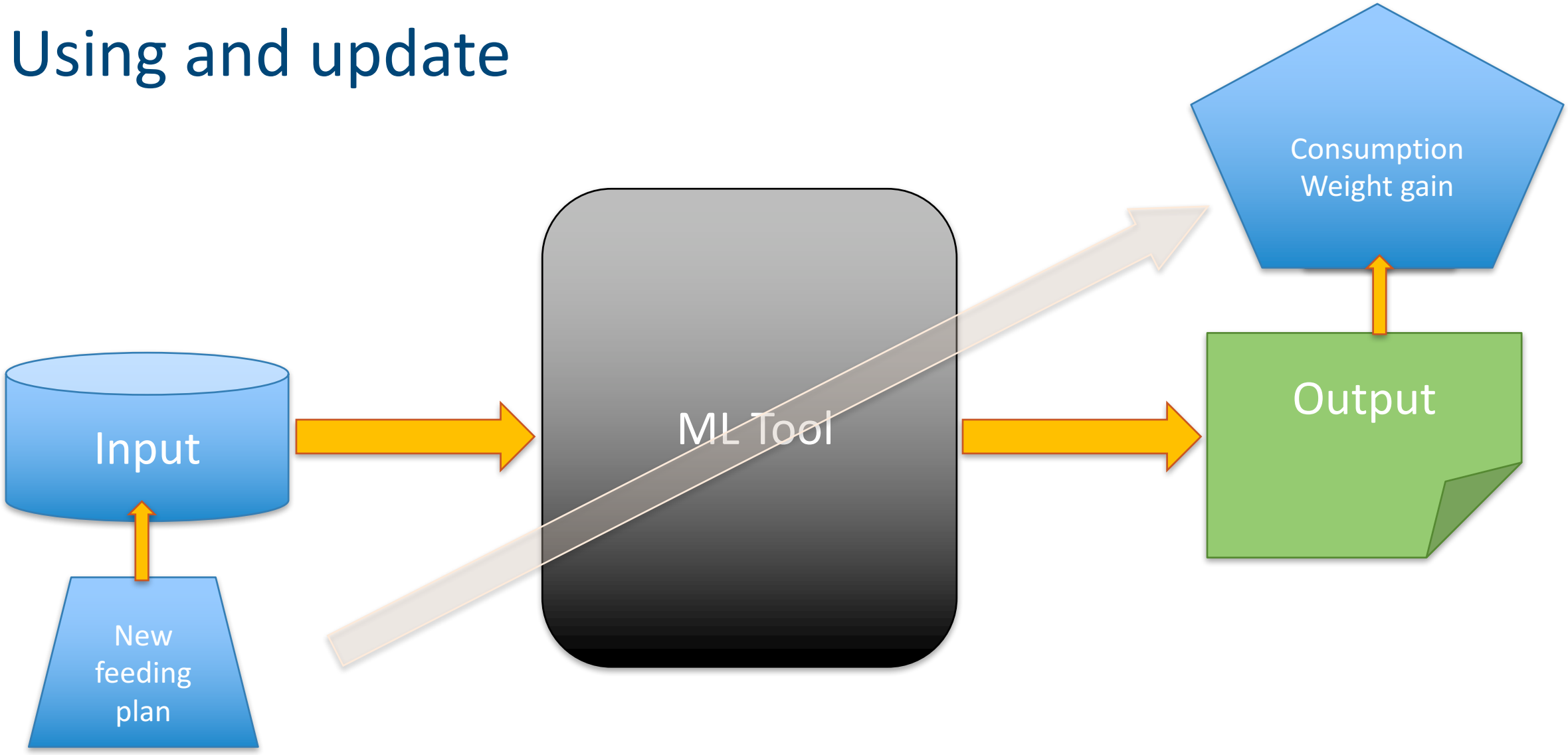


17

Using and update

18

Using and update



19

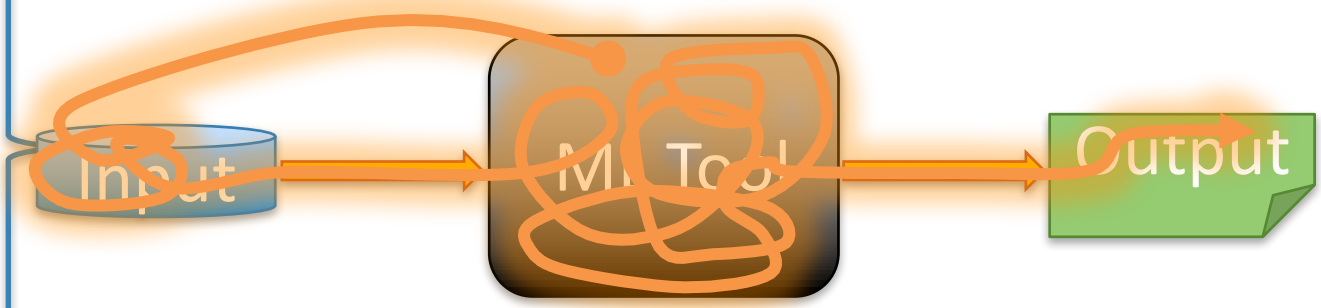
Using and update

New data :

Phase	1			2			3		
Période	1 → 16			17 → 22			23 → 43		
Proportion de lipides	20%			17%			18%		
Proportion de protéines	40%			45%			48%		
Consommation	2785	…	2834	1983	…	2013	1675	…	1369
Prise de poids	2078	…	2107	1631	…	1960	965	…	941

Phase	1			2			3		
Période	1 → 16			17 → 22			23 → 43		
Proportion de lipides	23%			18%			20%		
Proportion de protéines	36%			47%			46%		
Consommation	2786	…	2827	1890	…	1976	1665	…	1678
Prise de poids	2059	…	2113	1434	…	1468	978	…	965

…



Update of the forecast tool

20

Data heterogeneity

21

Data heterogeneity

Pb :

Phase	1			2			3		
Période	1 → 16			17 → 24			25 → 43		
Proportion de lipides	20%			17%			18%		
Proportion de protéines	40%			45%			48%		
Consommation	2785	…	2834	2383	…	2130	1570	…	1890
Prise de poids	2078	…	2107	1430	…	1468	987	…	965

Phase	1	2	3
Période	1 → 16	17 → 22	23 → 43
Proportion de lipides	20%	17%	18%
Proport…			
Consom…			
Prise de…			

Phase	1	2	3
Période	1 → 16	17 → 22	23 → 43
Proportion de lipides	23%	18%	20%
Proport…			
Consom…			
Prise de…			

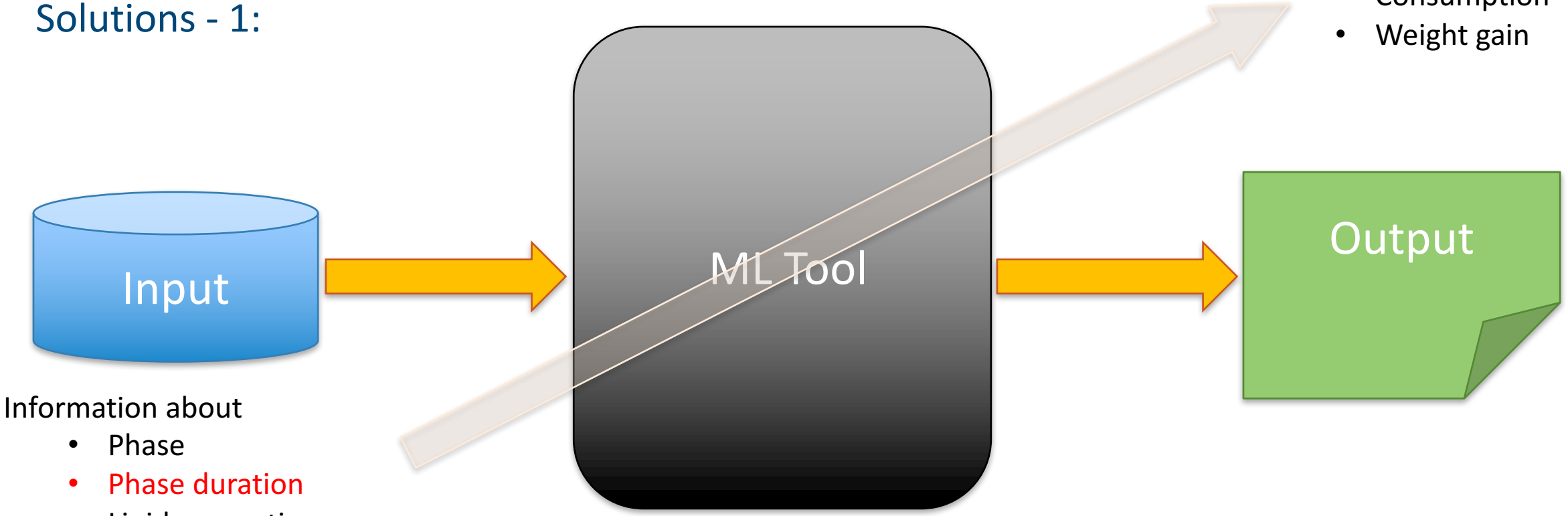
Phase	1			2			3		
Période	1 → 16			17 → 22			23 → 43		
Proportion de lipides	22%			16%			19%		
Proportion de protéines	36%			47%			46%		
Consommation	2786	…	2827	1890	…	1976	1665	…	1678
Prise de poids	2059	…	2113	1434	…	1468	978	…	965

…

22

Data heterogeneity

Solutions - 1:



For it works :

Enough data in each ”Phase duration” category

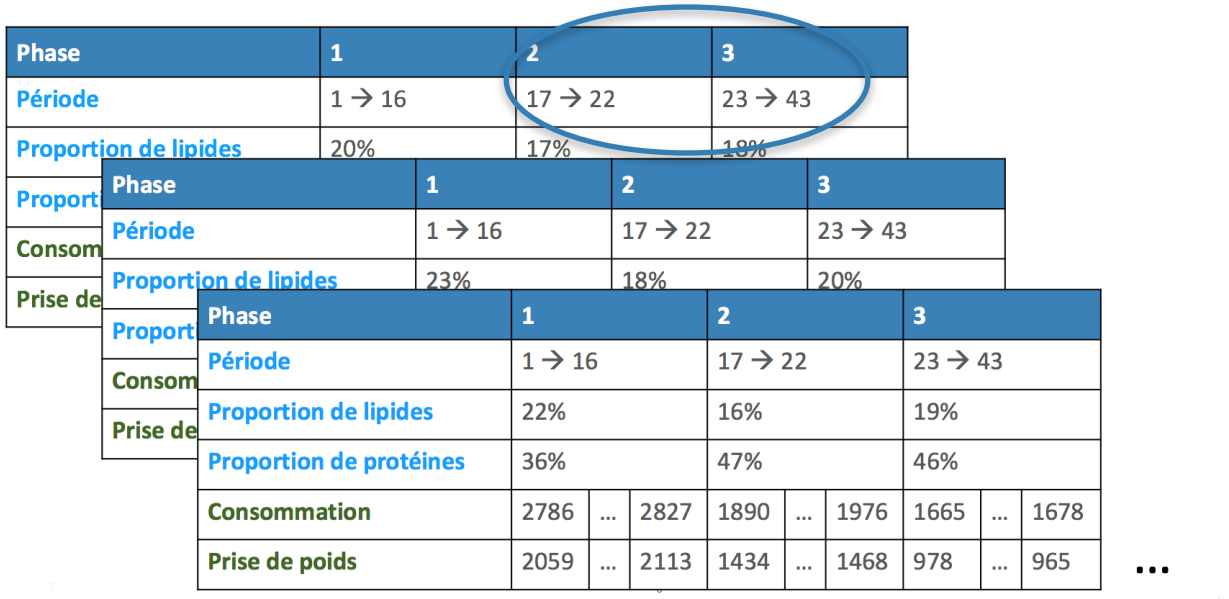
In our problem : not the case

23

Data heterogeneity

Solutions - 2:

Phase	1			2			3		
Période	1 → 16			17 → 24			25 → 43		
Proportion de lipides	20%			17%			18%		
Proportion de protéines	40%			45%			48%		
Consommation	2785	…	2834	1983	…	2013	1675	…	1369
Prise de poids	2078	…	2107	1631	…	1960	965	…	941



…

24

Data heterogeneity

Solutions - 2:

Phase	1			2			3		
Période	1 → 16			17 → 24			25 → 43		
Proportion de lipides	20%			17%			18%		
Proportion de protéines	40%			45%			48%		
Consommation	2785	…	2834	1983	…	2013	1675	…	1369
Prise de poids	2078	…	2107	1631	…	1960	965	…	941

Phase	1			2			3		
Période	1 → 16			17 → 22			23 → 43		
Proportion de lipides	20%			17%			18%		
Proportion de protéines	40%			45%			48%		
Consommation	2785	…	2834	2383	…	2130	1570	…	1890
Prise de poids	2078	…	2107	1430	…	1468	987	…	965

25

Data heterogeneity

Solutions - 2:

Phase	1			2			3		
Période	1 → 16			17 → 24			25 → 43		
Proportion de lipides	20%			17%			18%		
Proportion de protéines	40%			45%			48%		
Consommation	2785	...	2834	1983	...	2013	1675	...	1369
Prise de poids	2078	...	2107	1631	...	1960	965	...	941

Phase	1			2			3		
Période	1 → 16			17 → 22			23 → 43		
Proportion de lipides	20%			17%			18%		
Proportion de protéines	40%			45%			48%		
Consommation	2785	...	2834	2383	...	2130	1570	...	1890
Prise de poids	2078	...	2107	1430	...	1468	987	...	965

26

Data heterogeneity

Solutions - 2:

Update of the forecast tool

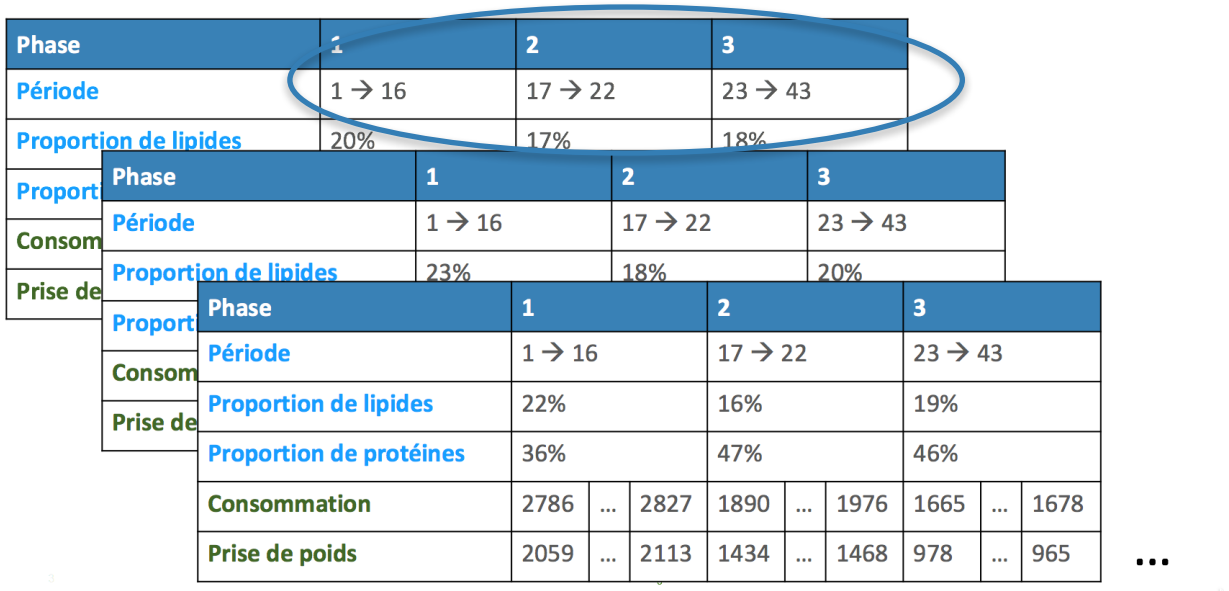
Phase	1	2	3
Période	1 → 16	17 → 22	23 → 43
Proportion de lipides	20%	17%	18%
Proportion de protéines	40%	45%	48%
Consommation	2785 … 2834	2383 … 2130	1570 … 1890
Prise de poids	2078 … 2107	1430 … 1468	987 … 965

27

Data heterogeneity

Limitations of implemented solution:

Phase	1			2			3		
Période	1 → 12			13 → 27			28 → 43		
Proportion de lipides	20%			17%			18%		
Proportion de protéines	40%			45%			48%		
Consommation	1785	…	1834	2983	…	2730	1170	…	1490
Prise de poids	1078	…	1107	1930	…	2368	887	…	865



1. Strong distortion :

Doubt about the solution accuracy

28

Data heterogeneity

Limitations of implemented solution:

Phase	1	2	3	4
Période	1 → 12	13 → 21	22 → 30	21 → 47
Proportion de lipides	20%	17%	18%	22%
Proportion de protéines	40%	45%	48%	42%
Consommation	...	...	...	...
Prise de poids	...	...	...	...

Phase	1	2	3
Période	1 → 16	17 → 22	23 → 43
Proportion de lipides	20%	17%	18%

Phase	1	2	3
Période	1 → 16	17 → 22	23 → 43
Proportion de lipides	23%	18%	20%

Phase	1	2	3
Période	1 → 16	17 → 22	23 → 43
Proportion de lipides	22%	16%	19%
Proportion de protéines	36%	47%	46%
Consommation	2786 ... 2827	1890 ... 1976	1665 ... 1678
Prise de poids	2059 ... 2113	1434 ... 1468	978 ... 965

...

2. different phase number :

Solution non implementable

29

Data heterogeneity

What to do?

Use mathematical and/or statistical modeling

Coupling model-data

30

MODEL

Equation for weight and consumption evolution

$$\frac{dP}{dt}(t) = \left(\alpha\, Q(t) + \beta\, L(t)\right)\left(\frac{\eta\, P(t)}{L(t)}\right)\frac{P(t) - \gamma}{\gamma}$$

$$\frac{dC}{dt}(t) = \frac{\eta\, P(t)}{L(t)}$$

$P(t)$: Weight at time t

$C(t)$: Full (since the beginning) food consumption at time t

32

Equation for weight and consumption evolution

$$\frac{dP}{dt}(t) = \left(\alpha Q(t) + \beta L(t)\right)\left(\frac{\eta P(t)}{L(t)}\right)\frac{P(t) - \gamma}{\gamma}$$

$$\frac{dC}{dt}(t) = \frac{\eta P(t)}{L(t)}$$

Speed of weight gain at time t

Speed of full food increase at time t

Protein proportion in the food

Food daily consumption

Lipid proportion in the food

Growth limiter

33

Equation for weight and consumption evolution

$$\frac{dP}{dt}(t) = \left(\alpha Q(t) + \beta L(t)\right)\left(\frac{\eta P(t)}{L(t)}\right)\frac{P(t) - \gamma}{\gamma}$$

$$\frac{dC}{dt}(t) = \frac{\eta P(t)}{L(t)}$$

Parameters

34

It is an ODE system

$$\frac{dP}{dt}(t) = \left(\alpha\, Q(t) + \beta\, L(t)\right)\left(\frac{\eta\, P(t)}{L(t)}\right)\frac{P(t) - \gamma}{\gamma}$$

$$\frac{dC}{dt}(t) = \frac{\eta\, P(t)}{L(t)}$$

35

It is an ODE system

$$\frac{dP}{dt}(t) = \left(\alpha\, Q(t) + \beta\, L(t)\right)\left(\frac{\eta\, P(t)}{L(t)}\right)\frac{P(t) - \gamma}{\gamma}$$

$$\frac{dC}{dt}(t) = \frac{\eta\, P(t)}{L(t)}$$

$$C(t = 0) = 0$$

$$P(t = 0) = \text{Weight at beginning}$$

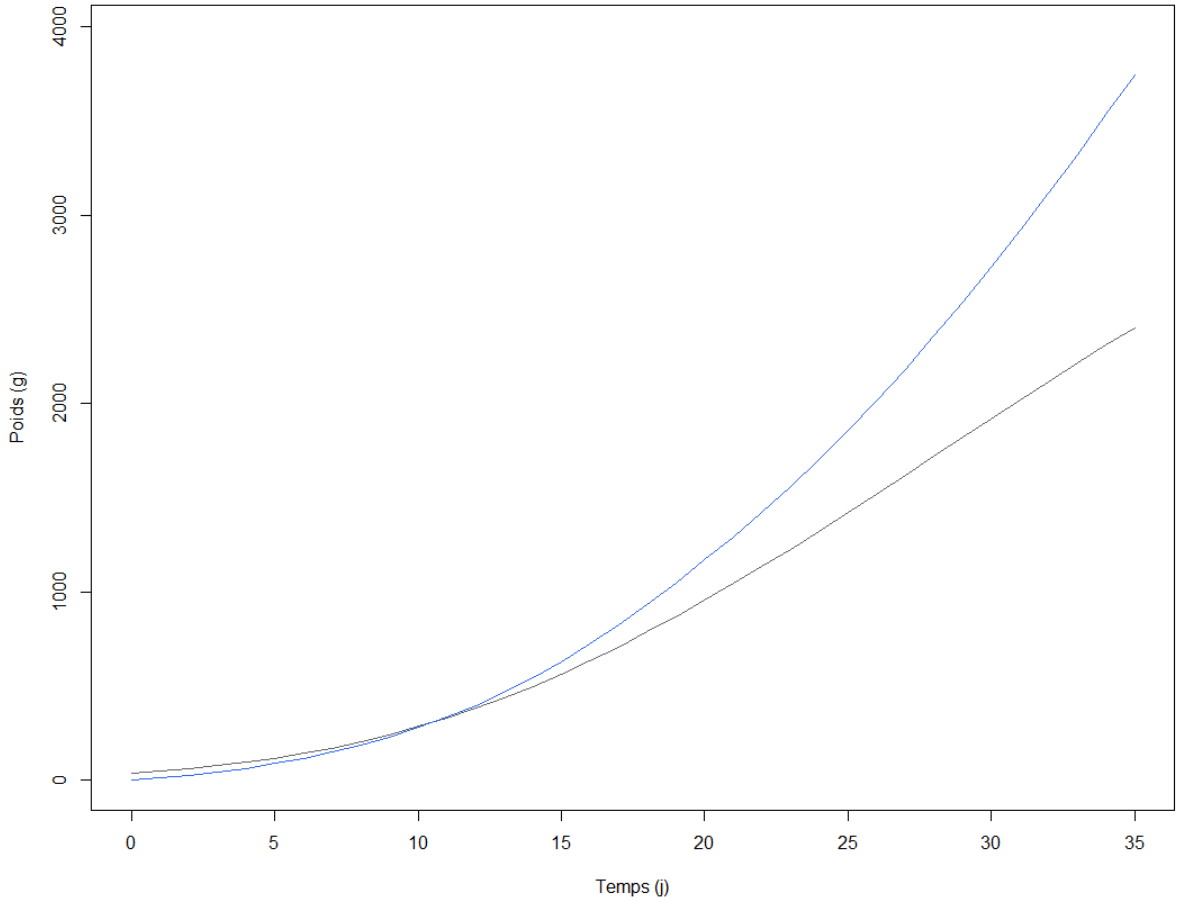
we can solve using R

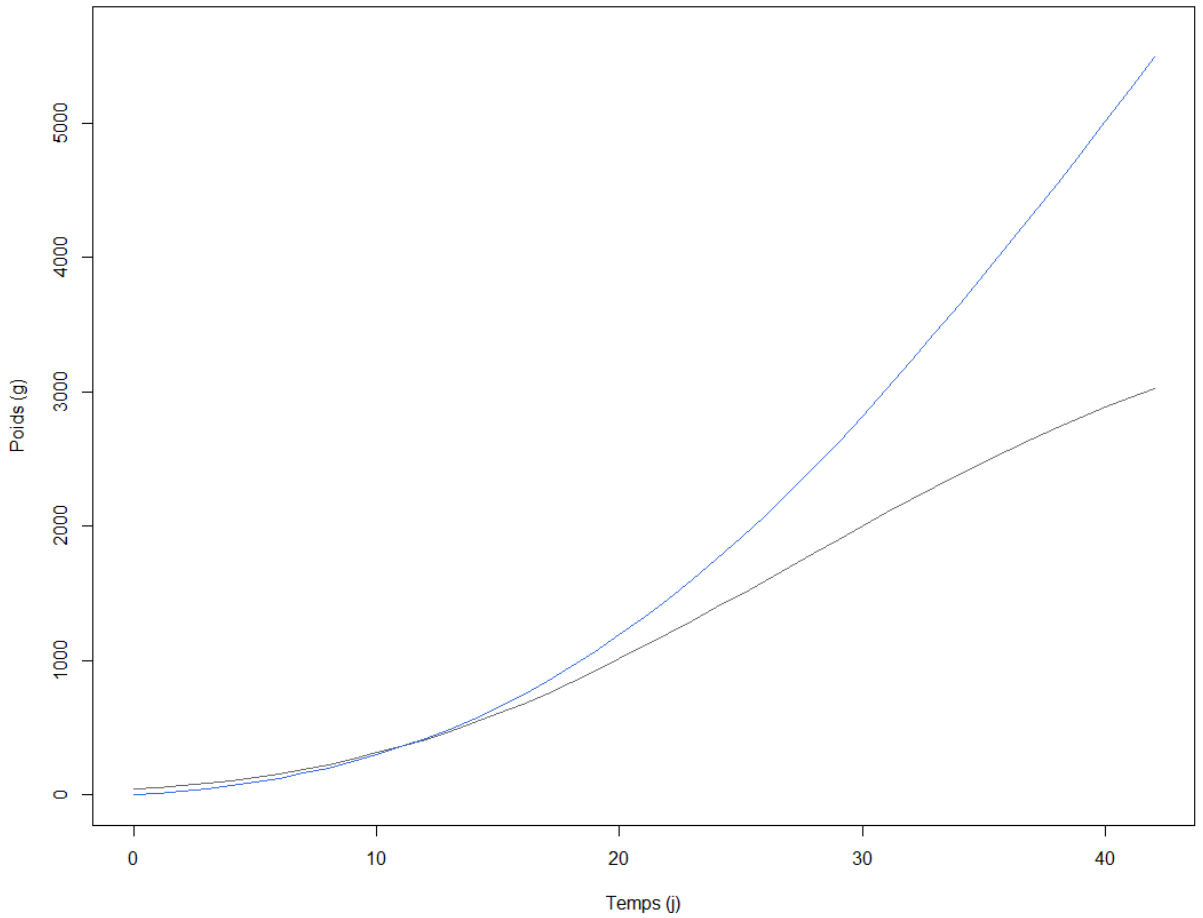
Matlab, Scilab, etc.

For given α, β, η and γ

36

Examples of solutions for given values of α, β, η, γ

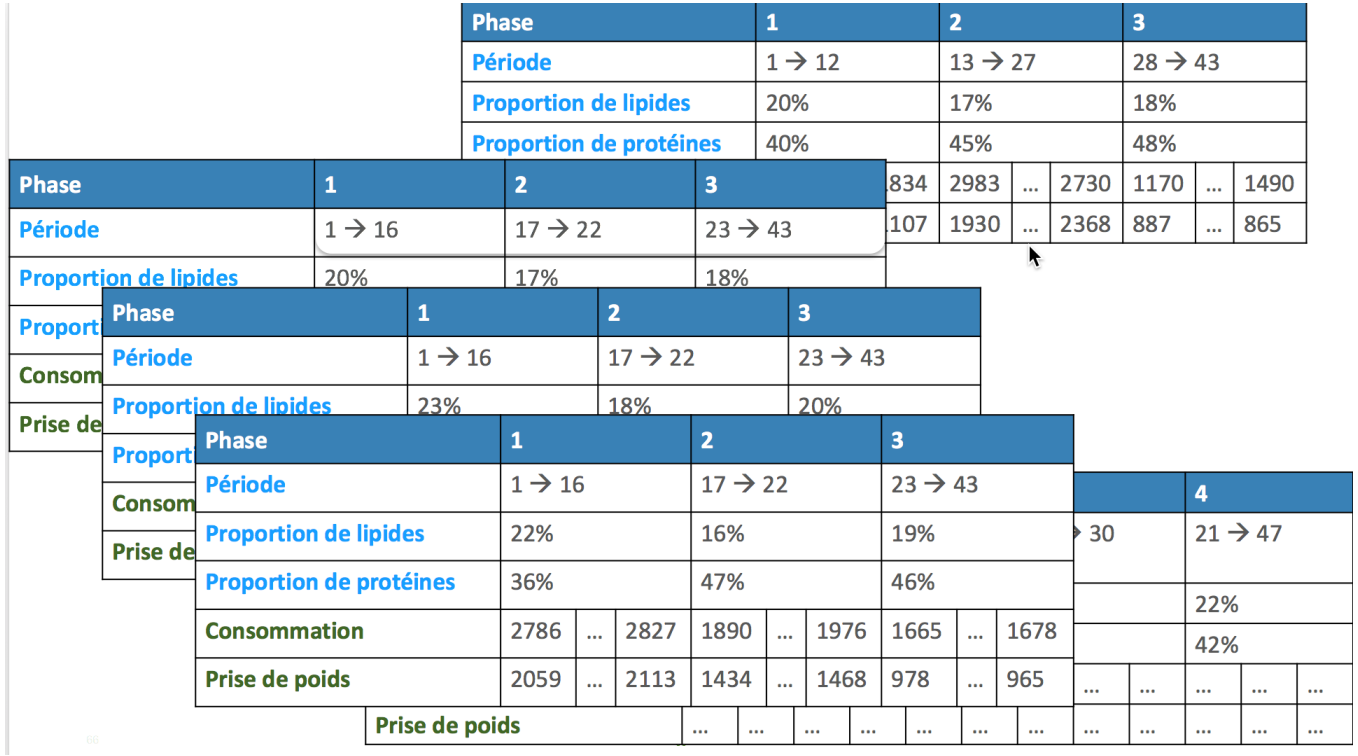




37

MODEL - DATA COUPLING

Interpretation of the data set according to the model



Phase	1	2	3
Période	1 → 12	13 → 27	28 → 43
Proportion de lipides	20%	17%	18%
Proportion de protéines	40%	45%	48%
Consommation	2834 ... 2827	2983 ... 2730	1170 ... 1490
Prise de poids	2107 ... 2113	1930 ... 2368	887 ... 865

Phase	1	2	3
Période	1 → 16	17 → 22	23 → 43
Proportion de lipides	20%	17%	18%
Proport			
Consom			
Prise de			

Phase	1	2	3
Période	1 → 16	17 → 22	23 → 43
Proportion de lipides	23%	18%	20%
Proport			
Consom			
Prise de			

Phase	1	2	3	4
Période	1 → 16	17 → 22	23 → 43	21 → 47
Proportion de lipides	22%	16%	19%	22%
Proportion de protéines	36%	47%	46%	42%
Consommation	2786 ... 2827	1890 ... 1976	1665 ... 1678	... ...
Prise de poids	2059 ... 2113	1434 ... 1468	978 ... 965	... ...

Prise de poids	... ...	... ...

39

Interpretation of the data set according to the model

Phase	1	2	3
Période	1 → 12	13 → 27	28 → 43
Proportion de lipides	20%	17%	18%
Proportion de protéines	40%	45%	48%
Consommation	2834 ... 2887	2983 ... 2730	1170 ... 1490
Prise de poids	2107 ... 2168	1930 ... 2368	887 ... 865

Phase	1	2	3
Période	1 → 16	17 → 22	23 → 43
Proportion de lipides	20%	17%	18%
Proport			
Consom			
Prise de			

Phase	1	2	3
Période	1 → 16	17 → 22	23 → 43
Proportion de lipides	23%	18%	20%
Proport			
Consom			
Prise de			

Phase	1	2	3	4
Période	1 → 16	17 → 22	23 → 43	21 → 47
Proportion de lipides	22%	16%	19%	22%
Proportion de protéines	36%	47%	46%	42%
Consommation	2786 ... 2827	1890 ... 1976	1665 ... 1678	... ...
Prise de poids	2059 ... 2113	1434 ... 1468	978 ... 965	... ...

Prise de poids	... ...	... ...

$$L(t), Q(t)$$

40

Interpretation of the data set according to the model

Phase	1	2	3
Période	1 → 12	13 → 27	28 → 43
Proportion de lipides	20%	17%	18%
Proportion de protéines	40%	45%	48%
Consommation	2834 ... 2983	... 2730	1170 ... 1490
Prise de poids	2107 ... 1930	... 2368	887 ... 865

Phase	1	2	3
Période	1 → 16	17 → 22	23 → 43
Proportion de lipides	20%	17%	18%
Proport			
Consom			
Prise de			

Phase	1	2	3
Période	1 → 16	17 → 22	23 → 43
Proportion de lipides	23%	18%	20%
Proport			
Consom			
Prise de			

Phase	1	2	3	4
Période	1 → 16	17 → 22	23 → 43	30 → 47
Proportion de lipides	22%	16%	19%	21 → 47
Proportion de protéines	36%	47%	46%	22%
Consommation	2786 ... 2827	1890 ... 1976	1665 ... 1678	42%
Prise de poids	2059 ... 2113	1434 ... 1468	978 ... 965	... ...
Prise de poids	... ...			... ...

$$L(t), Q(t)$$

$$\frac{dP}{dt}(t) = \left(\alpha\, Q(t) + \beta\, L(t)\right)\left(\frac{\eta\, P(t)}{L(t)}\right)\frac{P(t) - \gamma}{\gamma}$$

$$\frac{dC}{dt}(t) = \frac{\eta\, P(t)}{L(t)}$$

41

Interpretation of the data set according to the model

Phase	1	2	3
Période	1 → 12	13 → 27	28 → 43
Proportion de lipides	20%	17%	18%
Proportion de protéines	40%	45%	48%
Consommation	2834 ... 2983	... 2730	1170 ... 1490
Prise de poids	2107 ... 1930	... 2368	887 ... 865

Phase	1	2	3
Période	1 → 16	17 → 22	23 → 43
Proportion de lipides	20%	17%	18%
Proport			
Consom			
Prise de			

Phase	1	2	3
Période	1 → 16	17 → 22	23 → 43
Proportion de lipides	23%	18%	20%
Proport			
Consom			
Prise de			

Phase	1	2	3	4
Période	1 → 16	17 → 22	23 → 43	21 → 47
Proportion de lipides	22%	16%	19%	22%
Proportion de protéines	36%	47%	46%	42%
Consommation	2786 ... 2827	1890 ... 1976	1665 ... 1678	... ...
Prise de poids	2059 ... 2113	1434 ... 1468	978 ... 965	... ...
Prise de poids	... ...			... ...

$$L(t), Q(t)$$

$$\frac{dP}{dt}(t) = \left(\alpha\, Q(t) + \beta\, L(t)\right)\left(\frac{\eta\, P(t)}{L(t)}\right)\frac{P(t) - \gamma}{\gamma}$$

$$\frac{dC}{dt}(t) = \frac{\eta\, P(t)}{L(t)}$$

42

Phase	1	2	3
Période	1 → 12	13 → 27	28 → 43
Proportion de lipides	20%	17%	18%
Proportion de protéines	40%	45%	48%
Consommation	2834 ... 2983	... 2730	1170 ... 1490
Prise de poids	2107 ... 1930	... 2368	887 ... 865

Phase	1	2	3
Période	1 → 16	17 → 22	23 → 43
Proportion de lipides	20%	17%	18%
Proport			
Consom			
Prise de			

Phase	1	2	3
Période	1 → 16	17 → 22	23 → 43
Proportion de lipides	23%	18%	20%
Proport			
Consom			
Prise de			

Phase	1	2	3	4
Période	1 → 16	17 → 22	23 → 43	21 → 47
Proportion de lipides	22%	16%	19%	22%
Proportion de protéines	36%	47%	46%	42%
Consommation	2786 ... 2827	1890 ... 1976	1665 ... 1678	... ...
Prise de poids	2059 ... 2113	1434 ... 1468	978 ... 965	... ...
Prise de poids	... ...			... ...

$$L(t), Q(t)$$

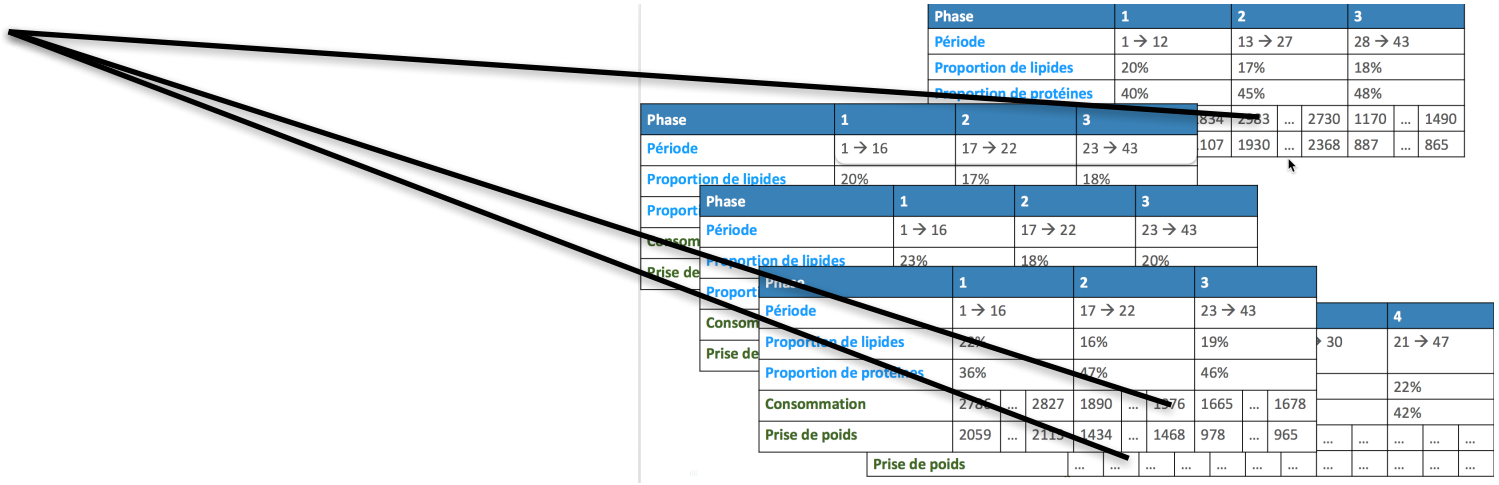
$$\frac{dP}{dt}(t) = \left(\alpha\, Q(t) + \beta\, L(t)\right)\left(\frac{\eta\, P(t)}{L(t)}\right)\frac{P(t) - \gamma}{\gamma}$$

$$\frac{dC}{dt}(t) = \frac{\eta\, P(t)}{L(t)}$$

43

Adjust α, β, η and γ

For each character (i)
and each value set of
α, β, η and γ

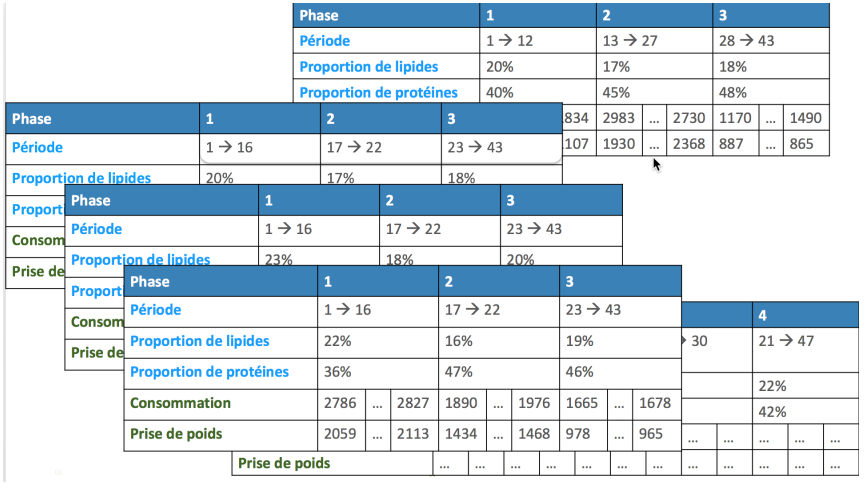


- Compute $P(t)$ et $C(t)$
- Compute P(End of phase 1), C(End of phase 1), P(End of phase 2), etc.
- Compute (P(End of phase 1) - Real Weight Gain(End of phase 1))^2 + (C(End of phase 1) - Real Consumption(End of phase 1))^2 + (P(End of phase 2) - Real Weight Gain(End of phase 2))^2 + … .
- = $D(i)$
- Compute $\sum_{all\ characters\ i} D(i)$
- = Fitness(α, β, η, γ)

44

Adjust α, β, η and γ

Using an optimization method:



Phase	1	2	3
Période	1 → 12	13 → 27	28 → 43
Proportion de lipides	20%	17%	18%
Proportion de protéines	40%	45%	48%

Phase	1	2	3
Période	1 → 16	17 → 22	23 → 43
Proportion de lipides	20%	17%	18%

Phase	1	2	3
Période	1 → 16	17 → 22	23 → 43
Proportion de lipides	23%	18%	20%

Phase	1	2	3
Période	1 → 16	17 → 22	23 → 43
Proportion de lipides	22%	16%	19%
Proportion de protéines	36%	47%	46%
Consommation	2786 ... 2827	1890 ... 1976	1665 ... 1678
Prise de poids	2059 ... 2113	1434 ... 1468	978 ... 965

																			4
																			21 → 47
																			22%
																			42%
Prise de poids	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

- Find $(\alpha, \beta, \eta, \gamma)$ minimizing :

$$\text{Fitness}(\alpha, \beta, \eta, \gamma) = \sum_{all\ characters\ i} D(i)$$

- Gives:

$$(\hat{\alpha}, \hat{\beta}, \hat{\eta}, \hat{\gamma})$$

45

What do we have ?

$$\frac{dP}{dt}(t) = \left(\hat{\alpha}\, Q(t) + \hat{\beta} L(t)\right) \left(\frac{\hat{\eta} P(t)}{L(t)}\right) \frac{P(t) - \hat{\gamma}}{\hat{\gamma}}$$

$$\frac{dC}{dt}(t) = \frac{\hat{\eta}\, P(t)}{L(t)}$$

$$C(t = 0) = 0$$

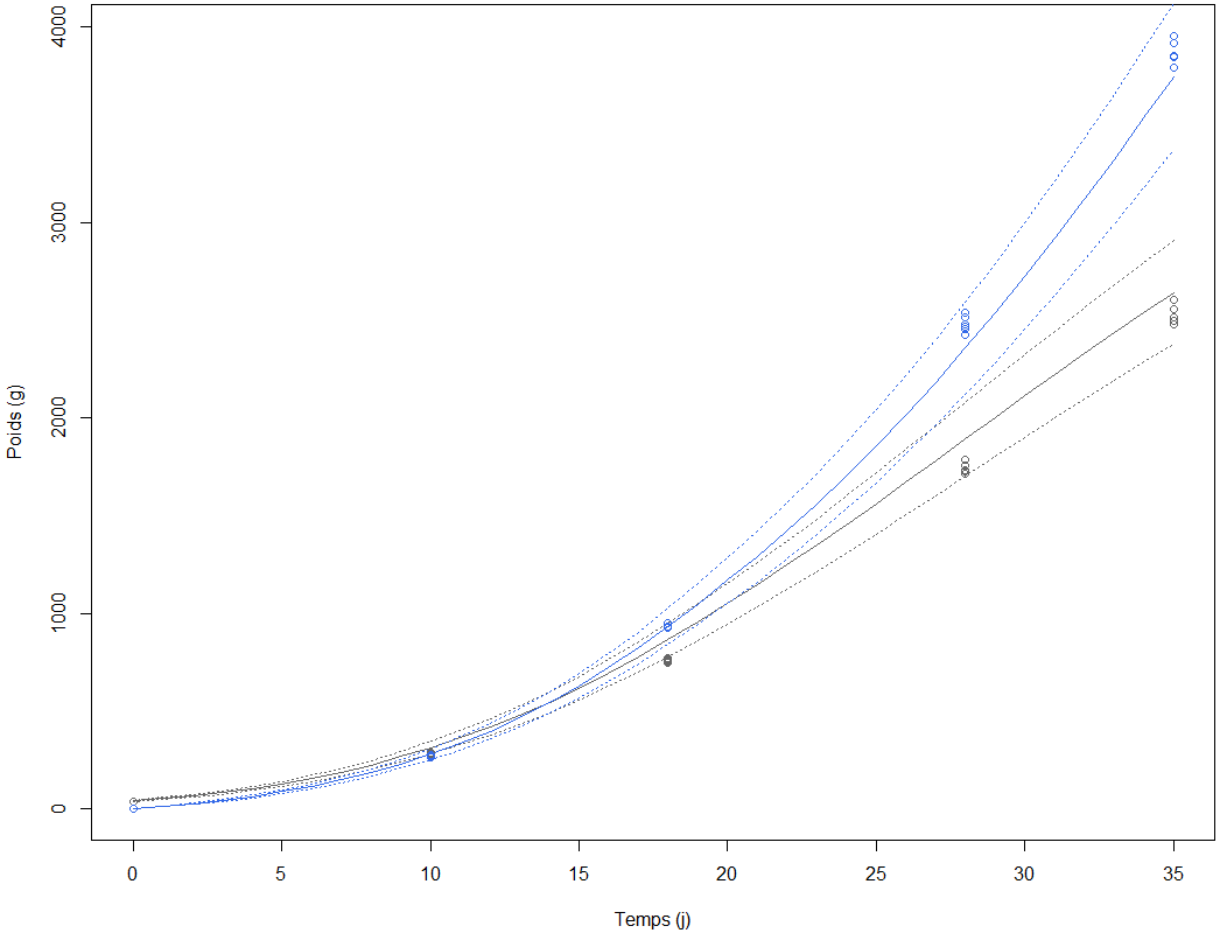
$$P(t = 0) = \text{Weight at beginning}$$

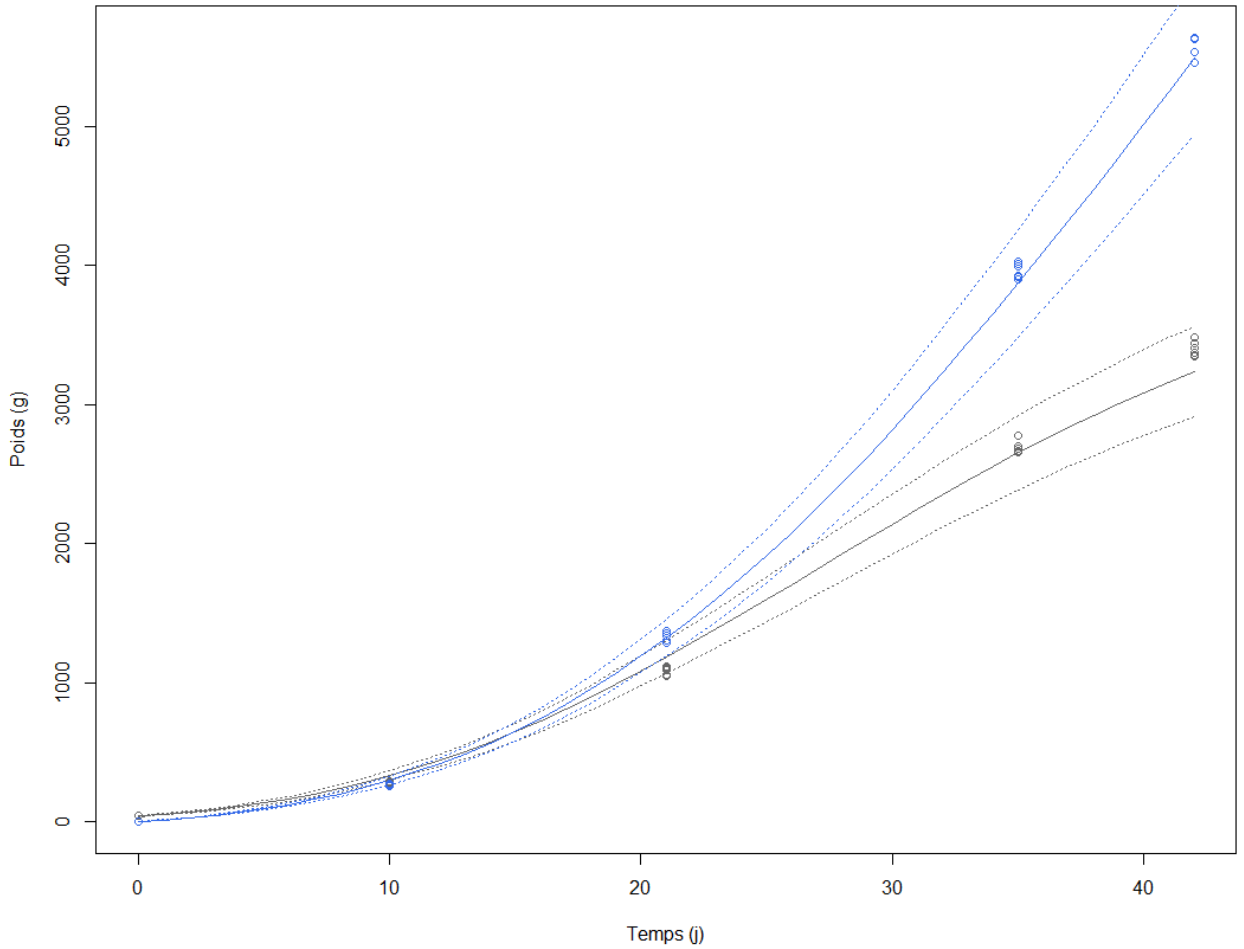
Allows computation of the weight and the full consumption at each time

46

What do we have ?

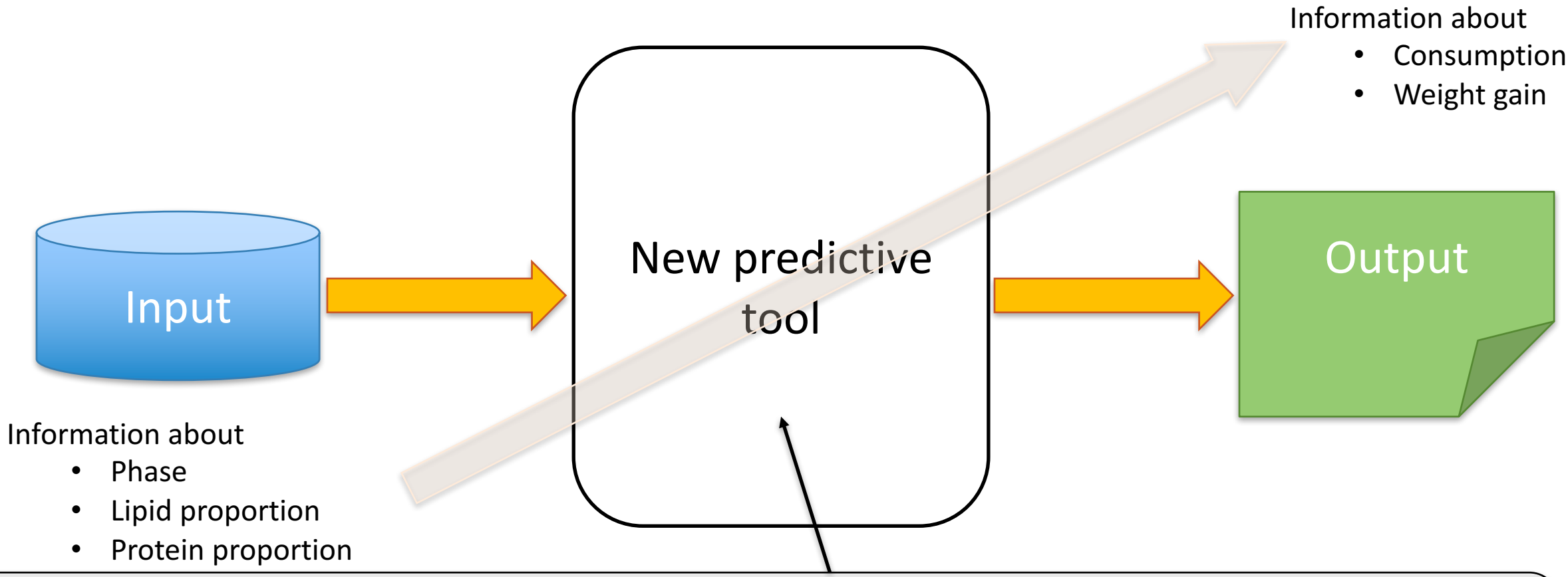
Allows computation of the weight and the full consumption at each time





47

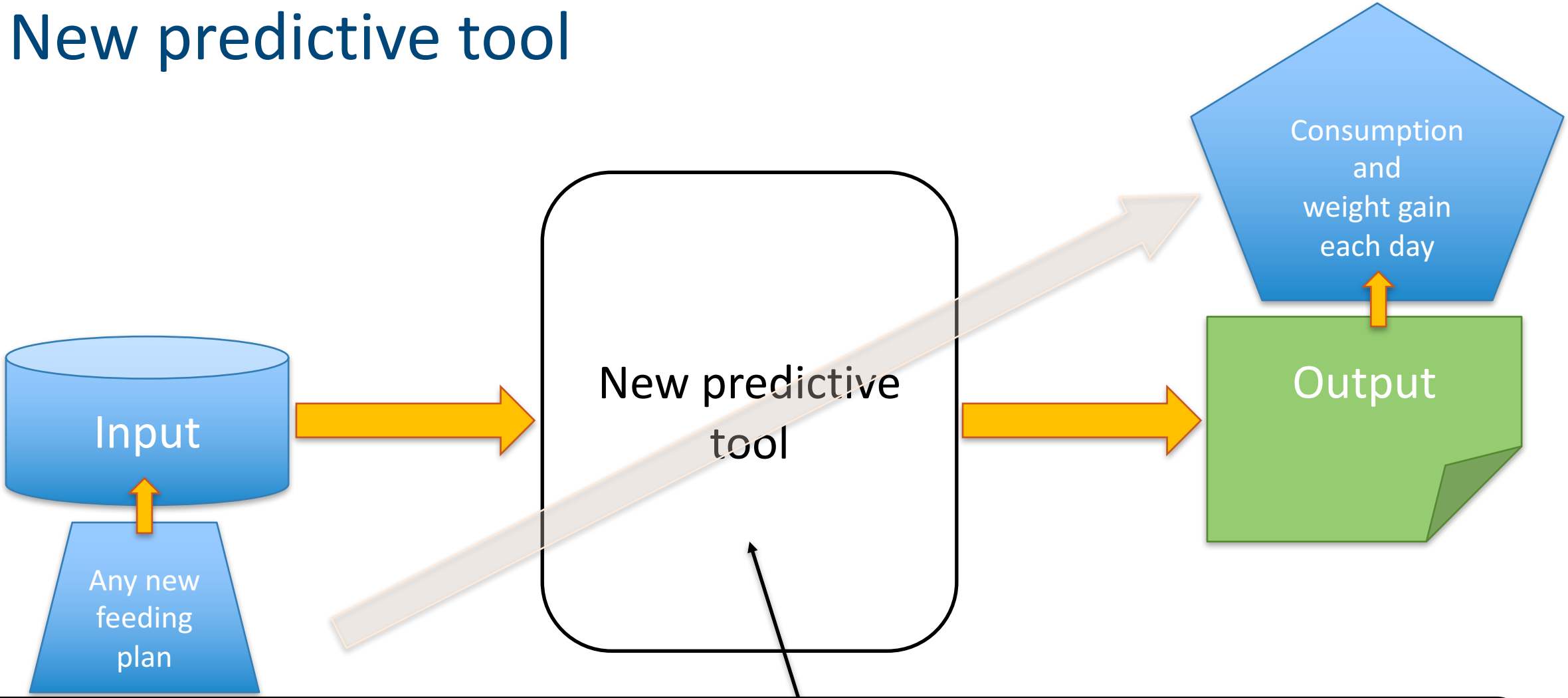
New predictive tool



$$\frac{dP}{dt}(t) = \left(\hat{\alpha}\, Q(t) + \hat{\beta} L(t)\right) \left(\frac{\hat{\eta} P(t)}{L(t)}\right) \frac{P(t) - \hat{\gamma}}{\hat{\gamma}} \qquad \frac{dC}{dt}(t) = \frac{\hat{\eta}\, P(t)}{L(t)}$$

48

New predictive tool



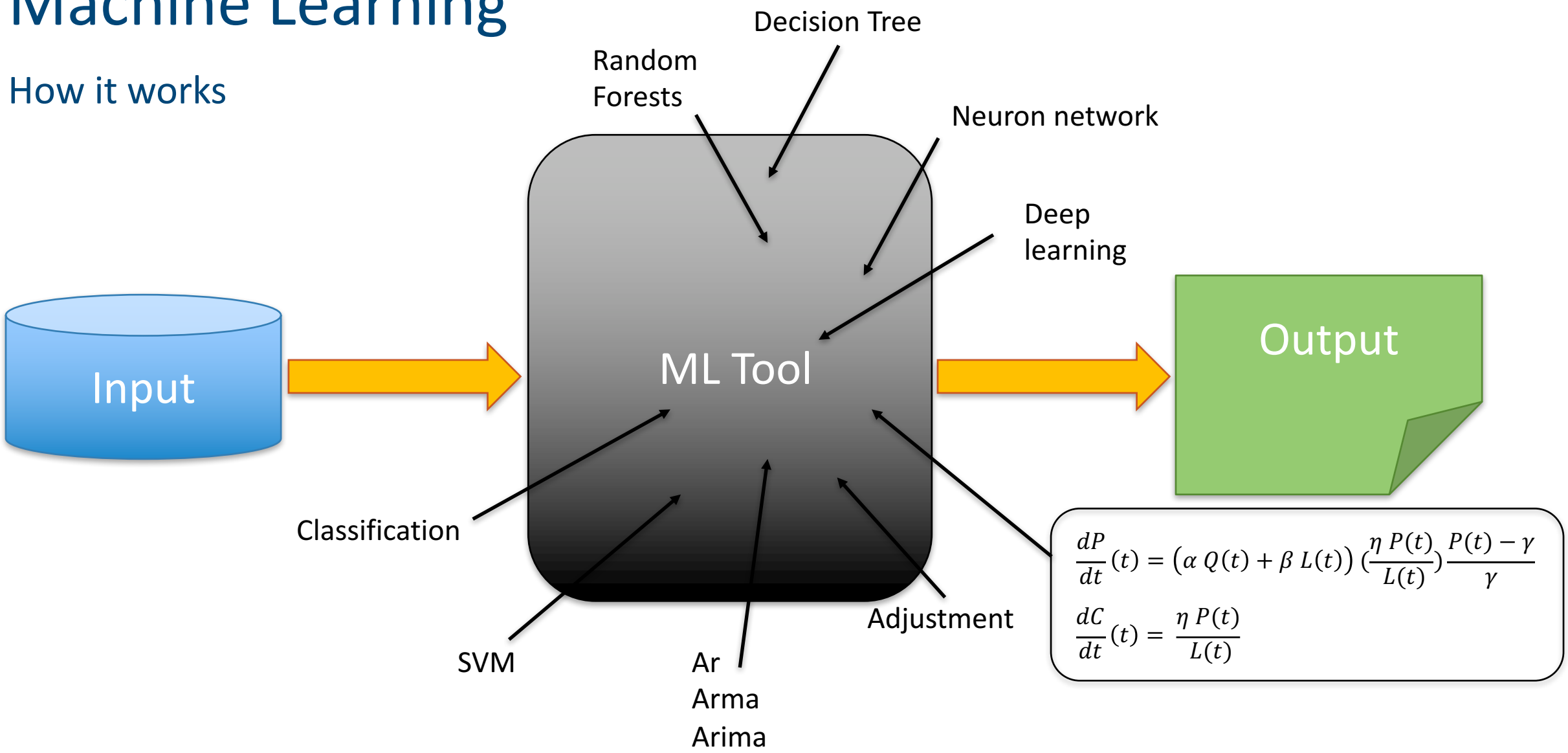
$$\frac{dP}{dt}(t) = \left(\hat{\alpha}\, Q(t) + \hat{\beta} L(t)\right) \left(\frac{\hat{\eta} P(t)}{L(t)}\right) \frac{P(t) - \hat{\gamma}}{\hat{\gamma}} \qquad \frac{dC}{dt}(t) = \frac{\hat{\eta}\, P(t)}{L(t)}$$

49

NEW MACHINE LEARNING TOOL BASED ON ODE DISCRETIZATION

Machine Learning

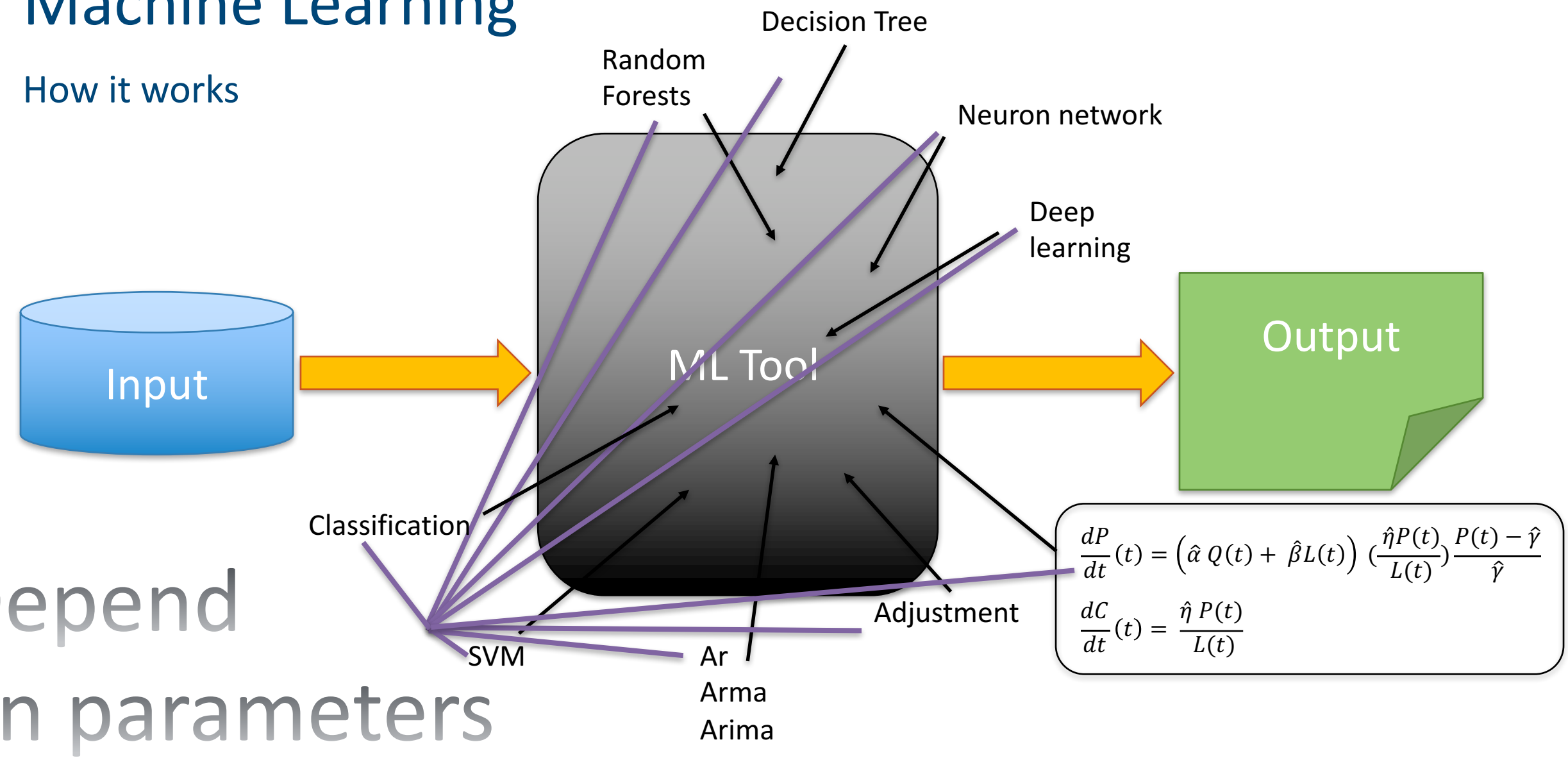
How it works



51

Machine Learning

How it works



Depend on parameters

52

Thanks for your attention