Open for Innovation
KNIME

The Creative Side of AI: Product Name Generation

Kathrin Melcher
kathrin.melcher@knime.com

© 2019 KNIME AG. All Right Reserved.

Agenda

- Neural network review
  - Feed-forward network
- Recurrent neural networks
  - What are recurrent neural networks?
  - Why do we need recurrent neural networks for sequential data?
  - What are LSTM units?
- Language modeling use case: Creating product names
  - Using the open source tool KNIME Analytics Platform

© 2019 KNIME AG. All Rights Reserved.

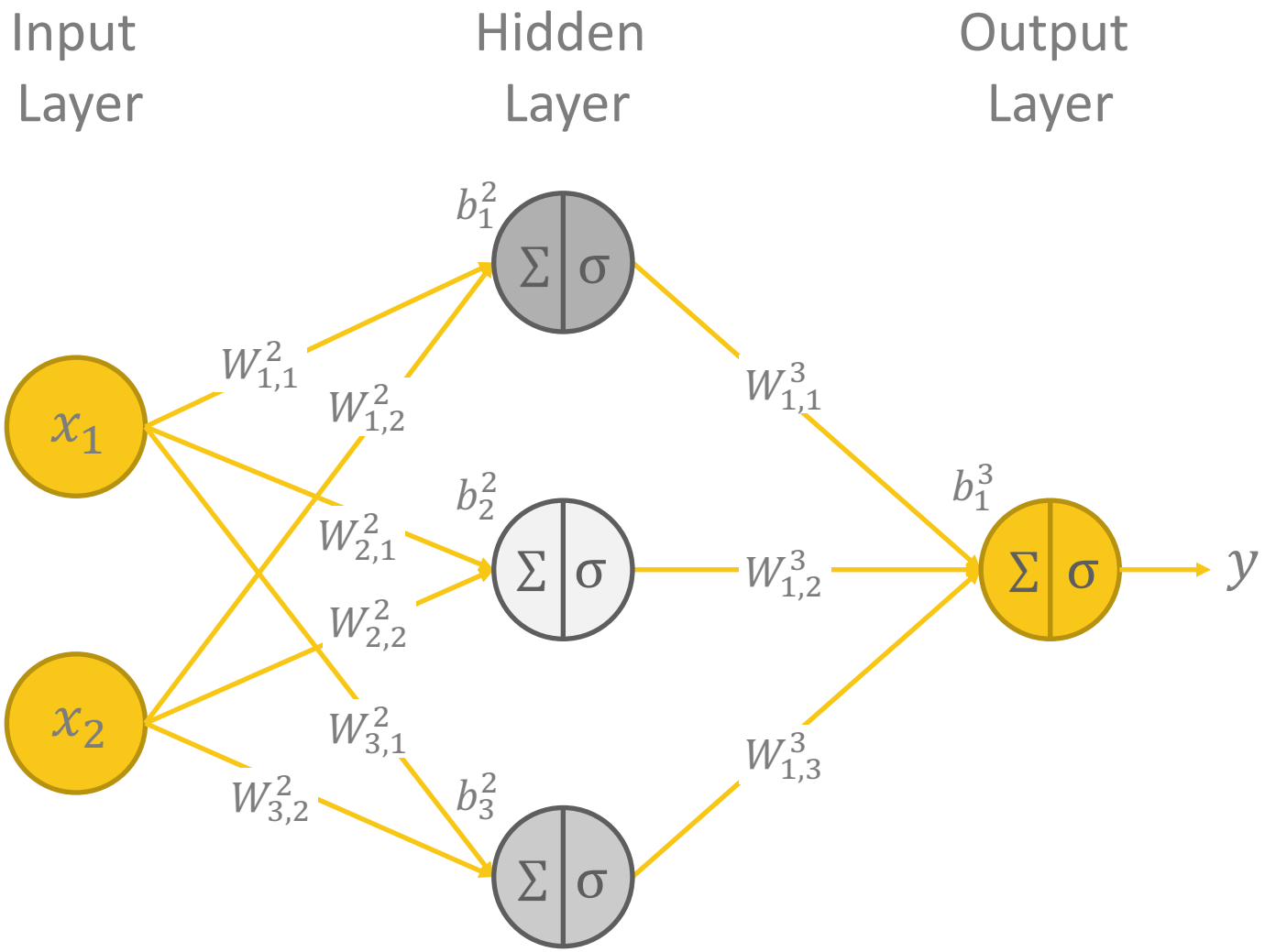
Open for Innovation ® KNIME

Neural Network Review

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ®
KNIME

Feed-forward networks



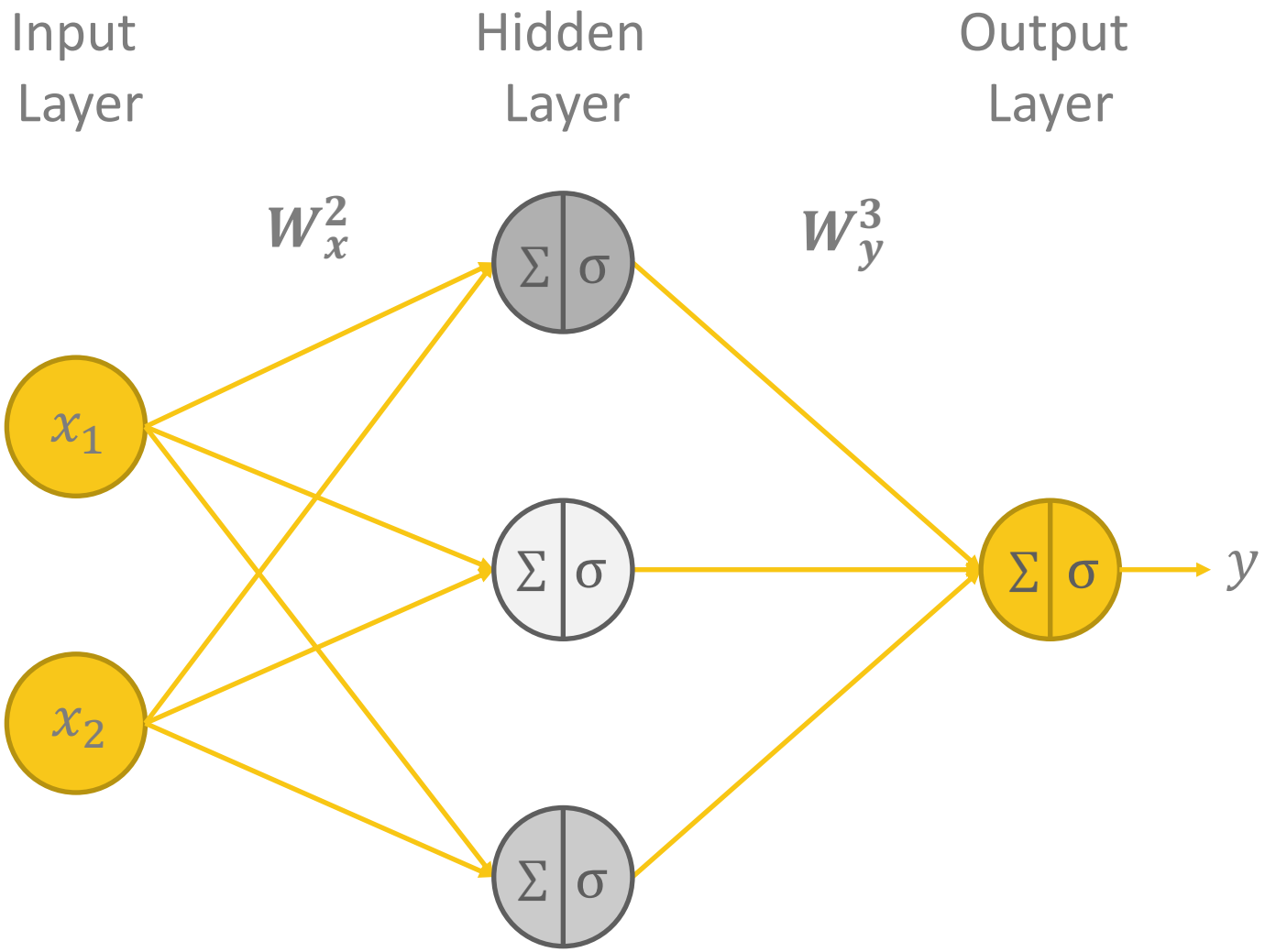
Forward pass:

$$a_2 = \sigma(W_x^2 x + b^2)$$
$$y = \sigma(W_y^3 a_2 + b^3)$$

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ® KNIME

Feed-forward networks - simplified



Forward pass:

$$a_2 = \sigma(W_x^2 x + b^2)$$
$$y = \sigma(W_y^3 a_2 + b^3)$$

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ® KNIME

What are RNNs and LSTMs?

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ®
KNIME

What are RNNs?

- Recurrent Neural Network (RNN) are a family of neural networks used for processing of sequential data
- RNNs are used for all sorts of tasks:
  - Language modeling
  - Text classification
  - Neural machine translation
  - Image captioning
  - Numerical time series data, e.g. sensor data

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ® KNIME

Why do we need RNNs for sequential data?

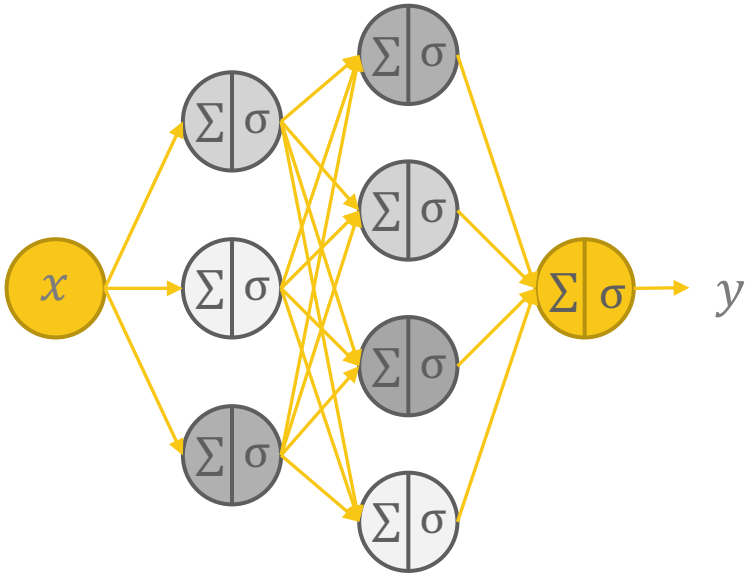
- Goal: Translation network from German to English

  “Ich mag Schokolade”
  => “I like chocolate”

- One option: Use feed forward network to translate word by word

- But what about this sentence?

  “Kathrin mag Schokolade”
  => “Kathrin likes chocolate”



Input x	Output y
Ich	I
mag	like
Schokolade	chocolate

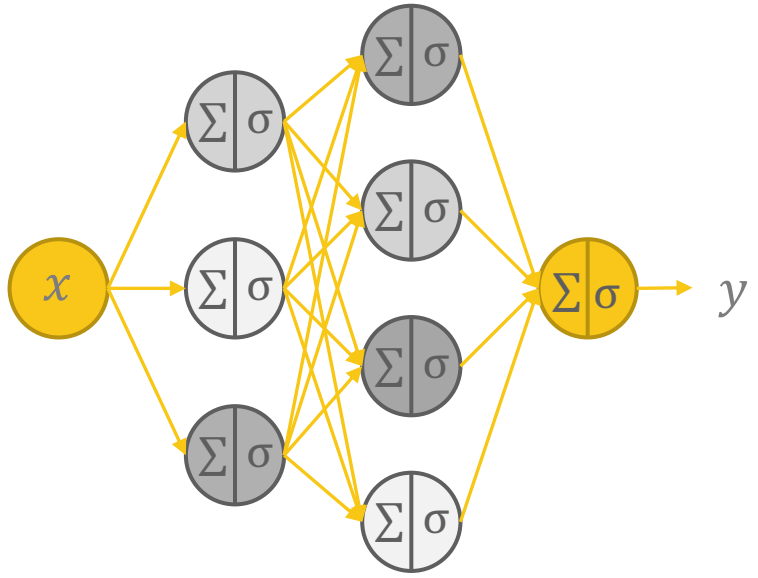
© 2019 KNIME AG. All Rights Reserved.

Open for Innovation KNIME

Why do we need RNNs for sequential data?

- Problems:
  - Each time step is completely independent
  - For translations we need context
  - More general: we need a network that remembers inputs from the past

- Solution: Recurrent neural networks



Input x	Output y
Ich	I
mag	like
Schokolade	chocolate

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation KNIME

What are RNNs?

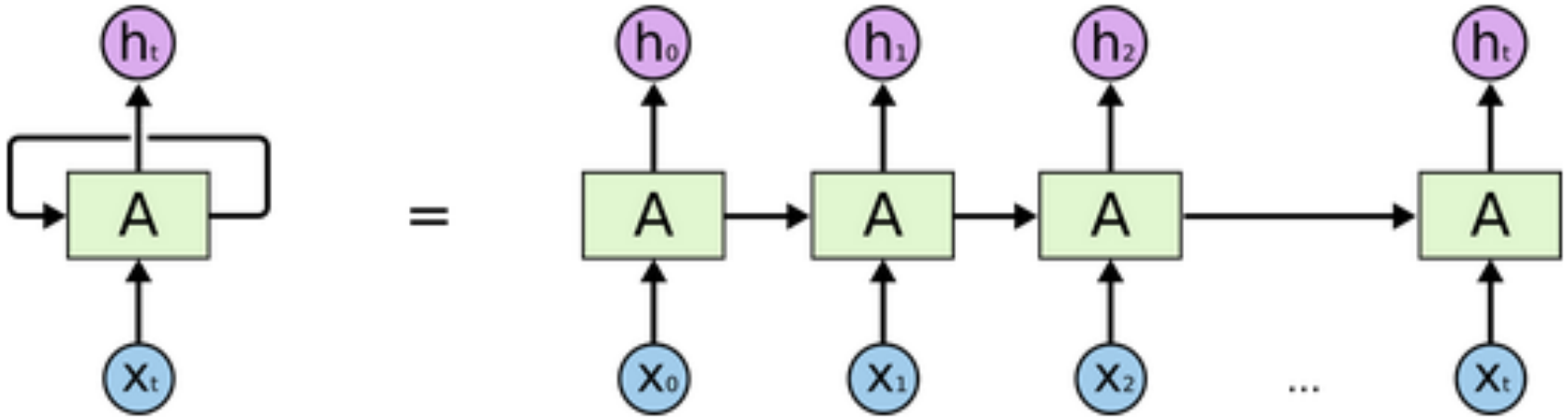
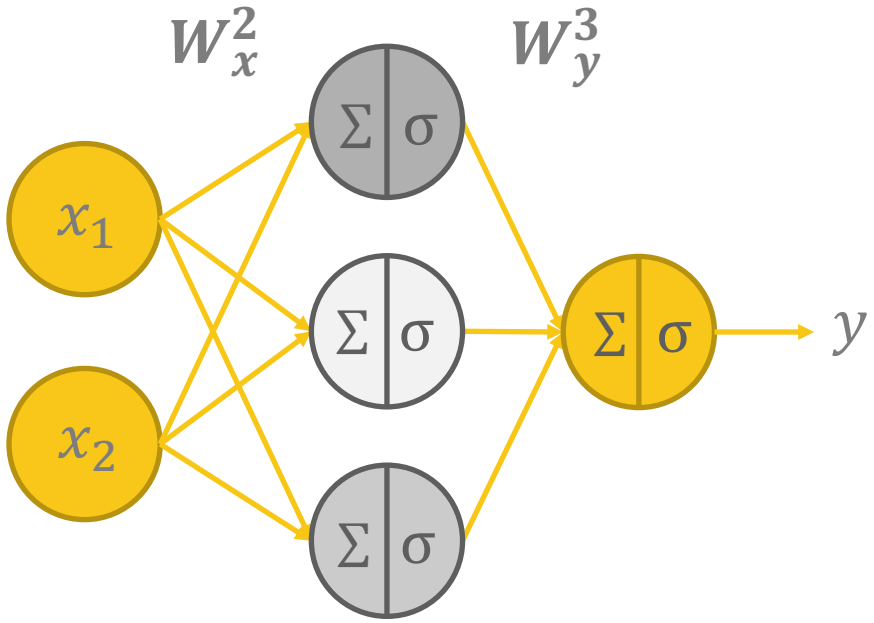


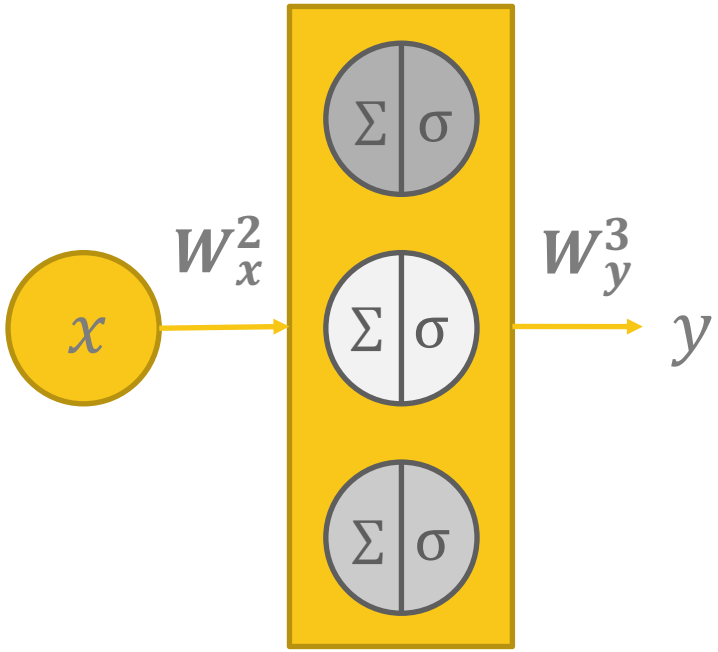
Image Source: Christopher Olah, https://colah.github.io/posts/2015-08-Understanding-LSTMs/

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ® KNIME

From feed forward to recurrent neural networks

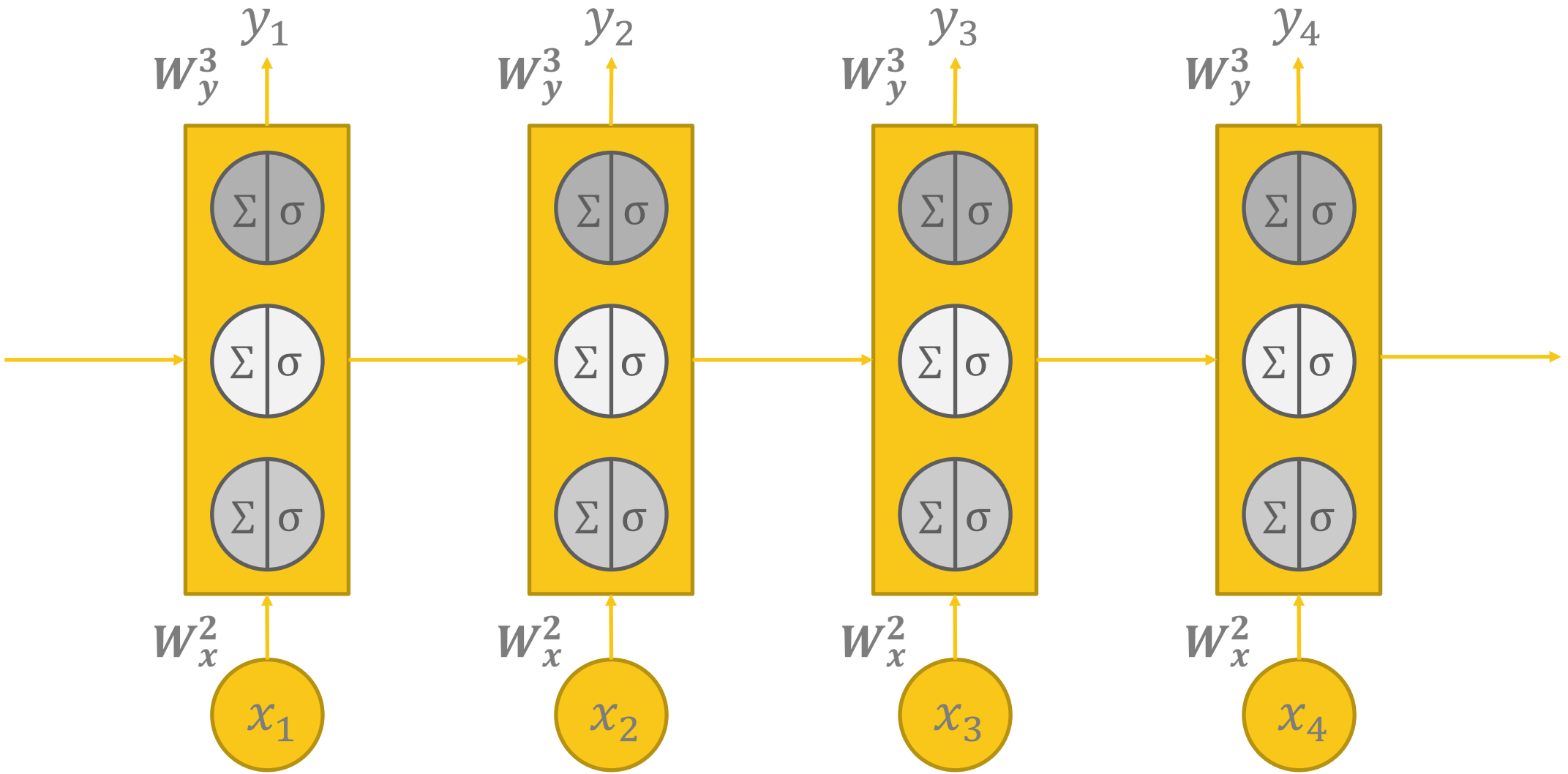




© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ® KNIME

From feed forward to recurrent neural networks



© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ® KNIME

Simple RNN unit

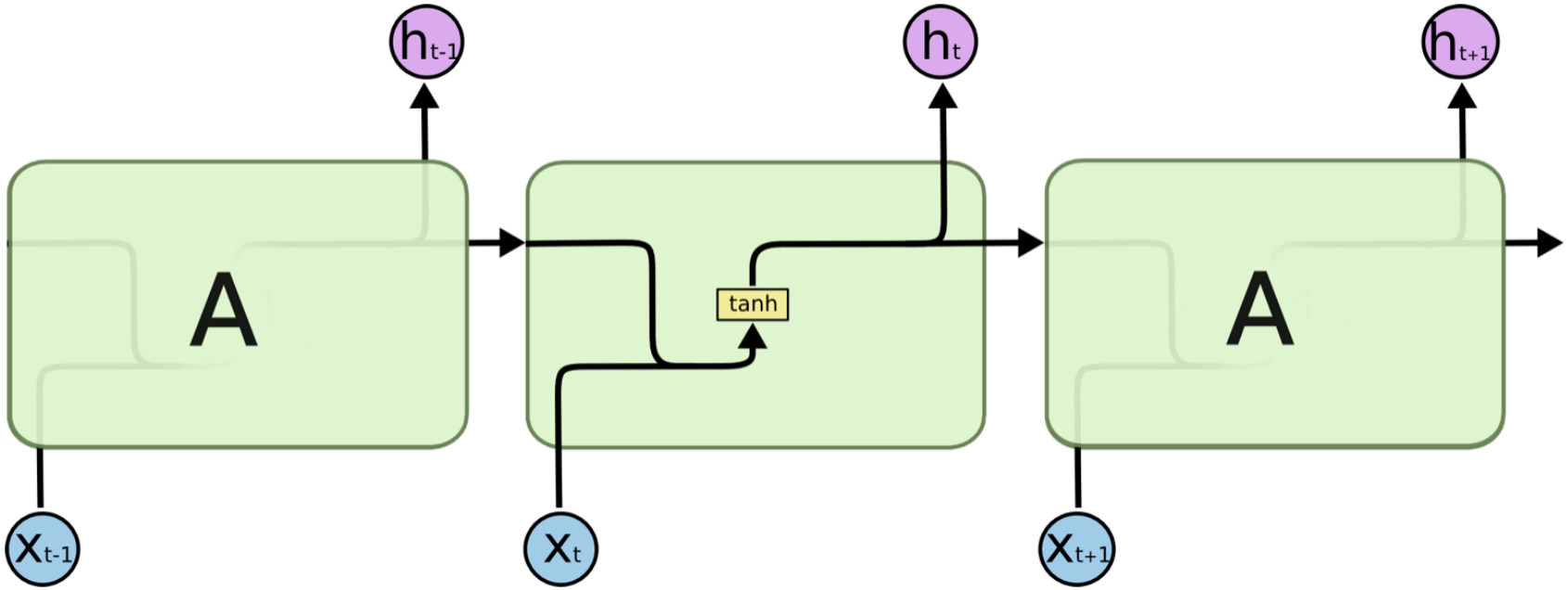


Image Source: Christopher Olah, https://colah.github.io/posts/2015-08-Understanding-LSTMs/

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ® KNIME

Limitations of simple layer structures

The “memory” of simple RNNs is sometimes too limited to be useful

- “Cars drive on the ___” (road)
- “I love the beach – my favorite sound is the crashing of the ____“ (cars? glass? waves?)

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ® KNIME

LSTM = Long Short Term Memory Unit

- Special type of unit with three gates
  - Forget gate
  - Input gate
  - Output gate

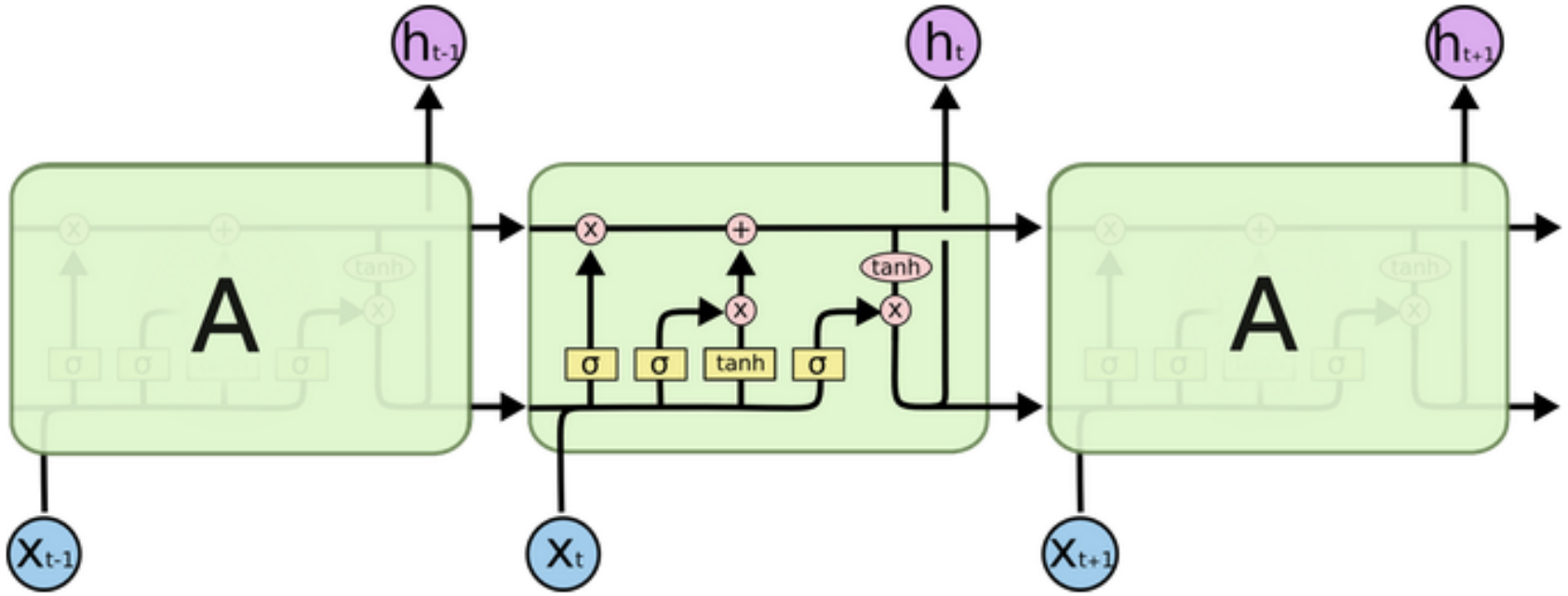


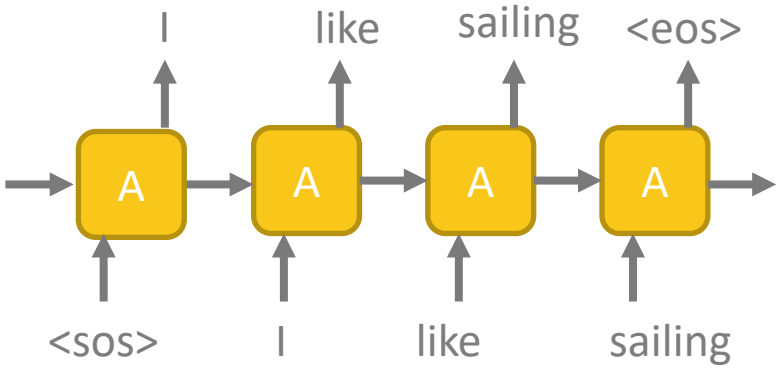
Image Source: Christopher Olah, https://colah.github.io/posts/2015-08-Understanding-LSTMs/

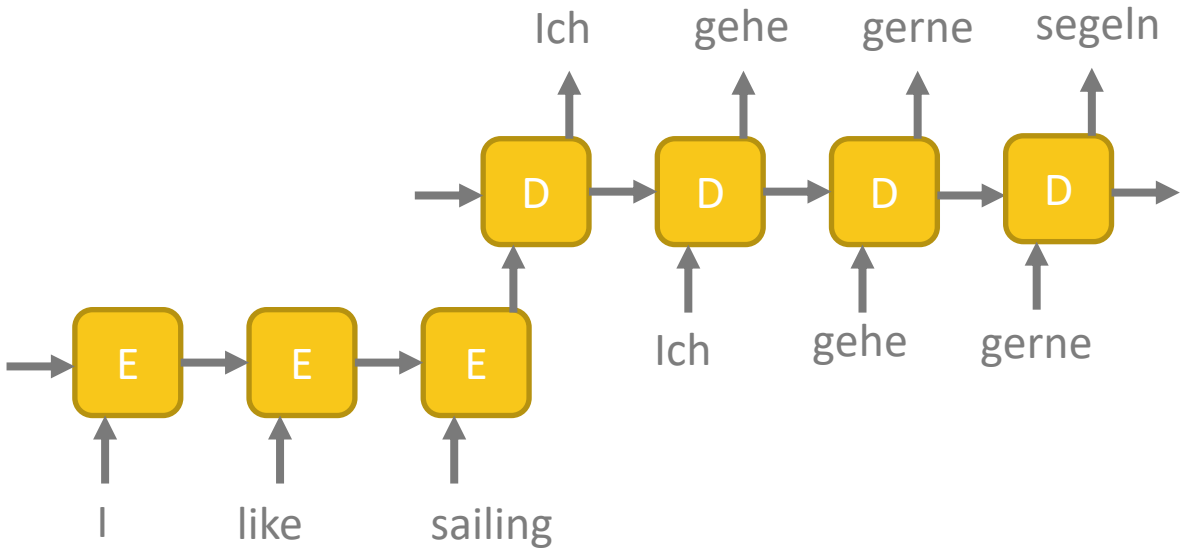
© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ® KNIME

Different network-structures and applications

Many to Many





Language model

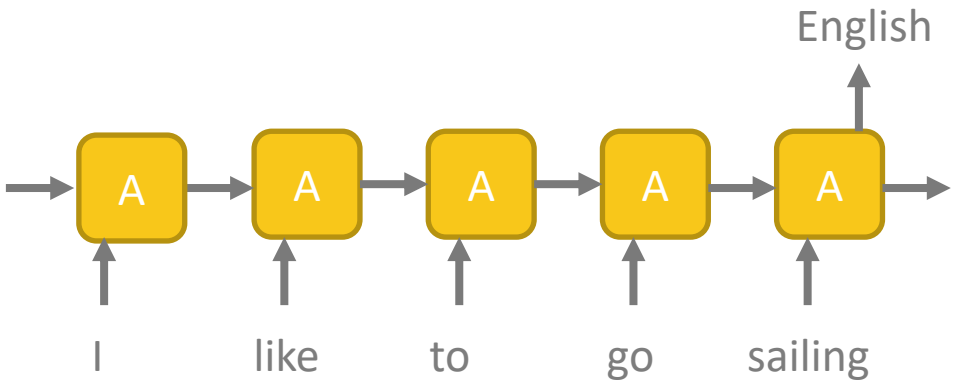
Neural machine translation

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ® KNIME

Different network-structures and applications

Many to one



Language classification

Text classification

One to many

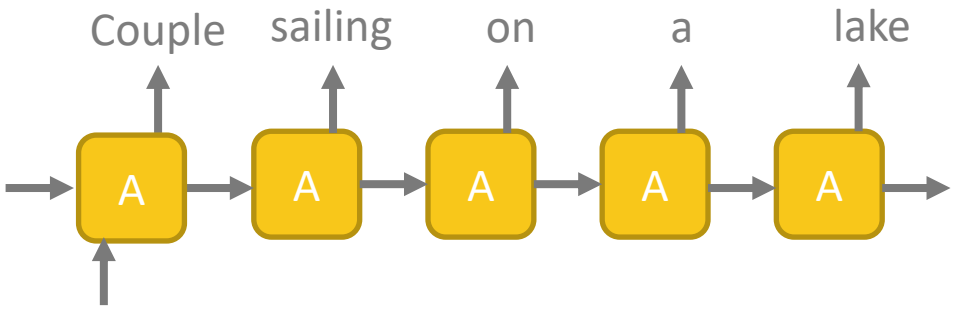


Image captioning

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ® KNIME

The Creative Side of AI: Product Name Generation

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ®
KNIME

Naming new products is not as easy at it might sound…

What if there was a way to help this process along, using neural network?

Product name requirements for our outdoor or hiking clothing line

- Needs to stand out
- Sound familiar to customers
- Evoke the feeling of nature
- Need many potential name to evaluate with the team

So what can we do?

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ® KNIME

Build a model to generate fictional mountain names

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ®
KNIME

Quick sneak preview – Which of these mountain names are real?

Set 1

Barlock Mountain
Casterland Mountain
Shafford Peak

Set 2

Roblin Hill
Buxley Mountain
Baldy Rock

Set 3

Terrey Hill
Walter Hill
Little Buck Butte

Correct answer: NONE!

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ® KNIME

The data set

Data source:

- 33,012 names of mountains in the US
- Extracted from Wikipedia via a Wikidata query

Row0	Telescope Peak
Row1	Half Dome
Row2	Crestone Needle
Row3	Mount Baker
Row4	Mount San Antonio
Row5	Mount Eisenhower
Row6	White Butte
Row7	Mount Moffett
Row8	Arrigetch Peaks
Row9	Belknap Crater
Row10	Mount Sheridan

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ® KNIME

The training data

Many to one

- Multiple training samples per name
- Define a window length, e.g. 3 for Mount Baker leads to:

Input sequence	Output
Mou	n
oun	t
unt	<space>
nt <space>	B

Many to many

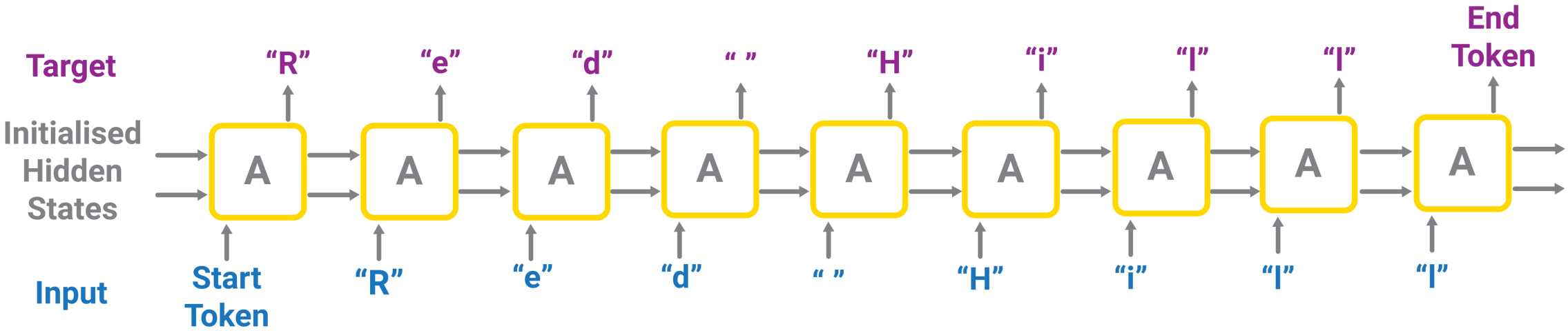
- One sample per name
- Shifted by one character

Input sequence	Output sequence
<S>Mount<space>Baker<E>	Mount<space>Baker<E>
<S>Red<space>Hill<E>	Red<space>Hill<E>
<S>Half<space>Dome<E>	Half<space>Dome<E>
<S>White<space>Butte<E>	White<space>Butte<E>

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ® KNIME

Many to many structure

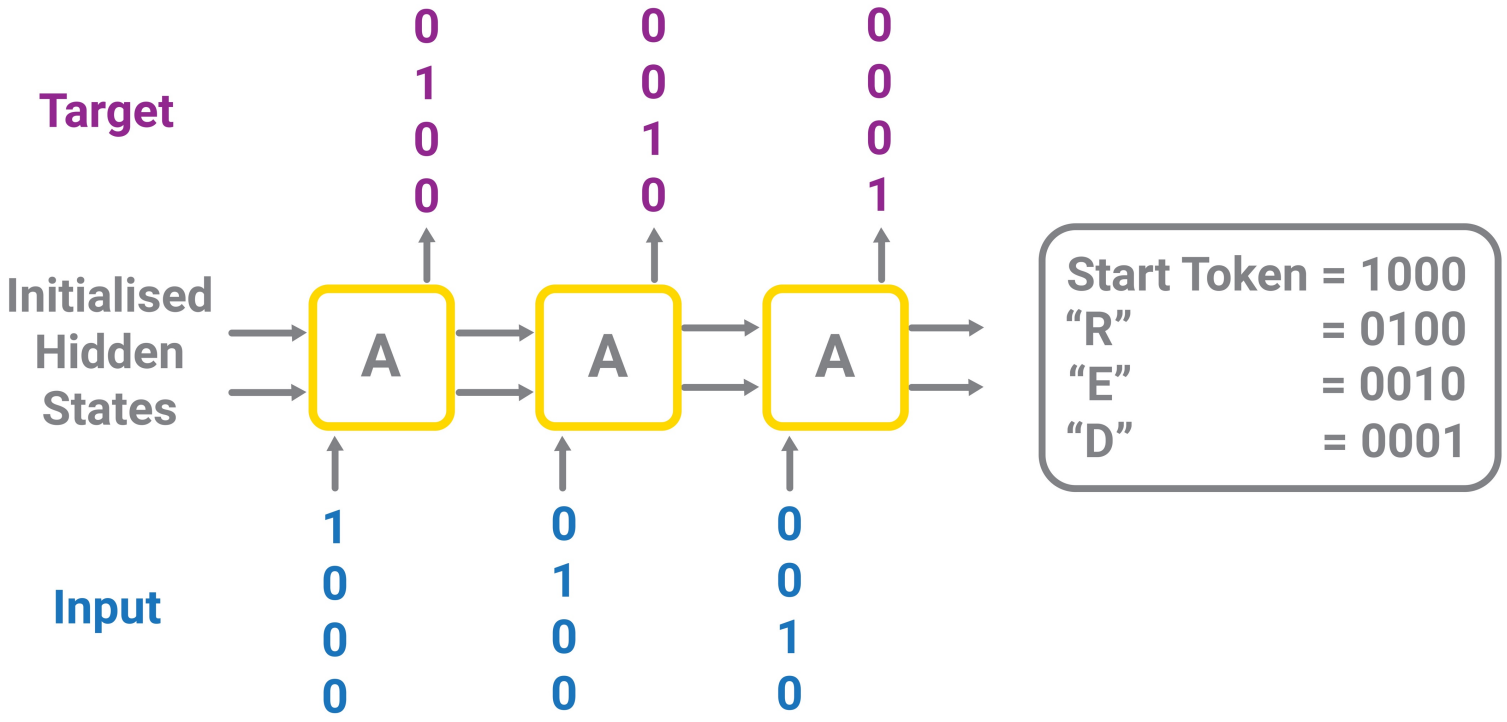


© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ® KNIME

Some data transformation

- Networks operate on numbers, not letters
- Represent each characters of our mountain names in a way that the model can understand in matrix space



© 2019 KNIME AG. All Rights Reserved.

Open for Innovation KNIME

Transformation steps

- Individual characters will be encoded as one-hot vectors
- We create a dictionary to transform names to vector sequences and vice versa
- Longest mountain name has 58 characters, so we pad out shorter names with zeros, since sequence lengths in each batch need to be the same

© 2019 KNIME AG. All Rights Reserved.

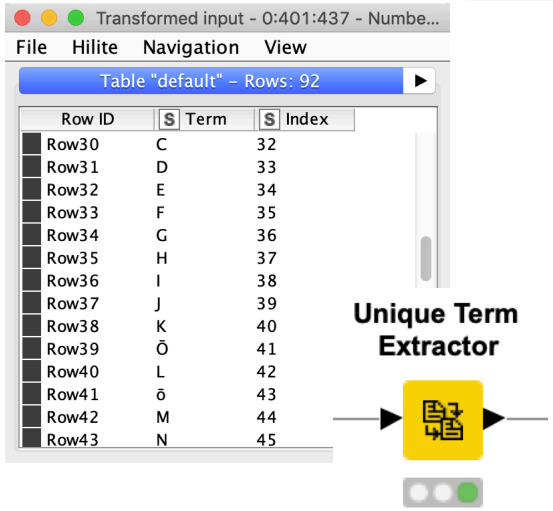
Open for Innovation ® KNIME

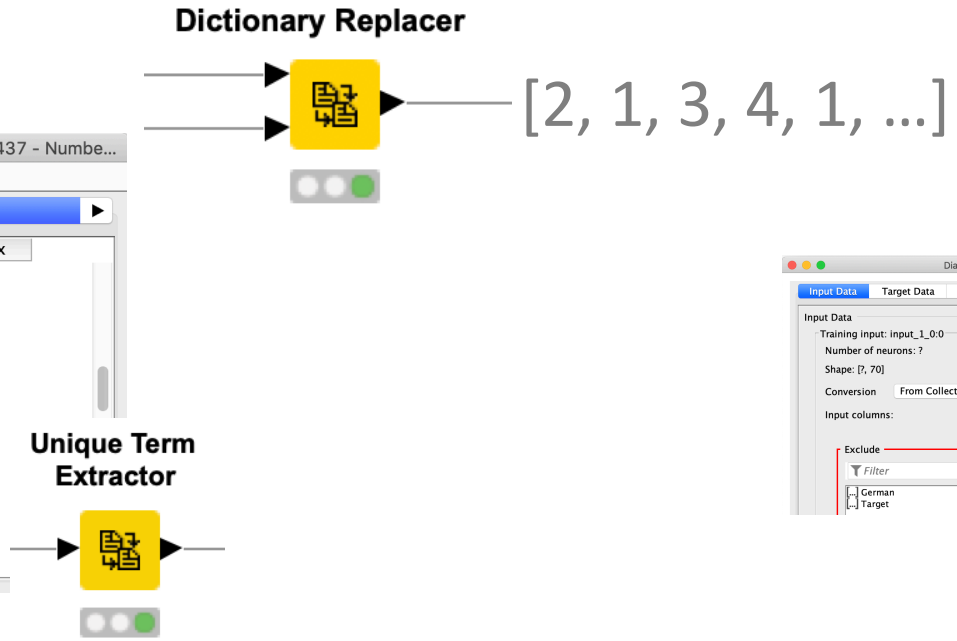
Two step encryption

Step 1: Index Based

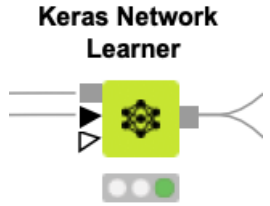
Step 2: One-Hot Encoding

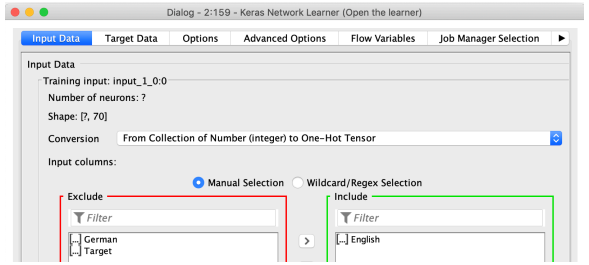
Red Hill





[2, 1, 3, 4, 1, …]





$$\begin{bmatrix}0\\1\\0\\0\\0\\0\end{bmatrix}\begin{bmatrix}1\\0\\0\\0\\0\\0\end{bmatrix}\begin{bmatrix}0\\0\\1\\0\\0\\0\end{bmatrix}\begin{bmatrix}0\\0\\0\\1\\0\\0\end{bmatrix}\begin{bmatrix}1\\0\\0\\0\\0\\0\end{bmatrix}$$

© 2019 KNIME AG. All Rights Reserved.

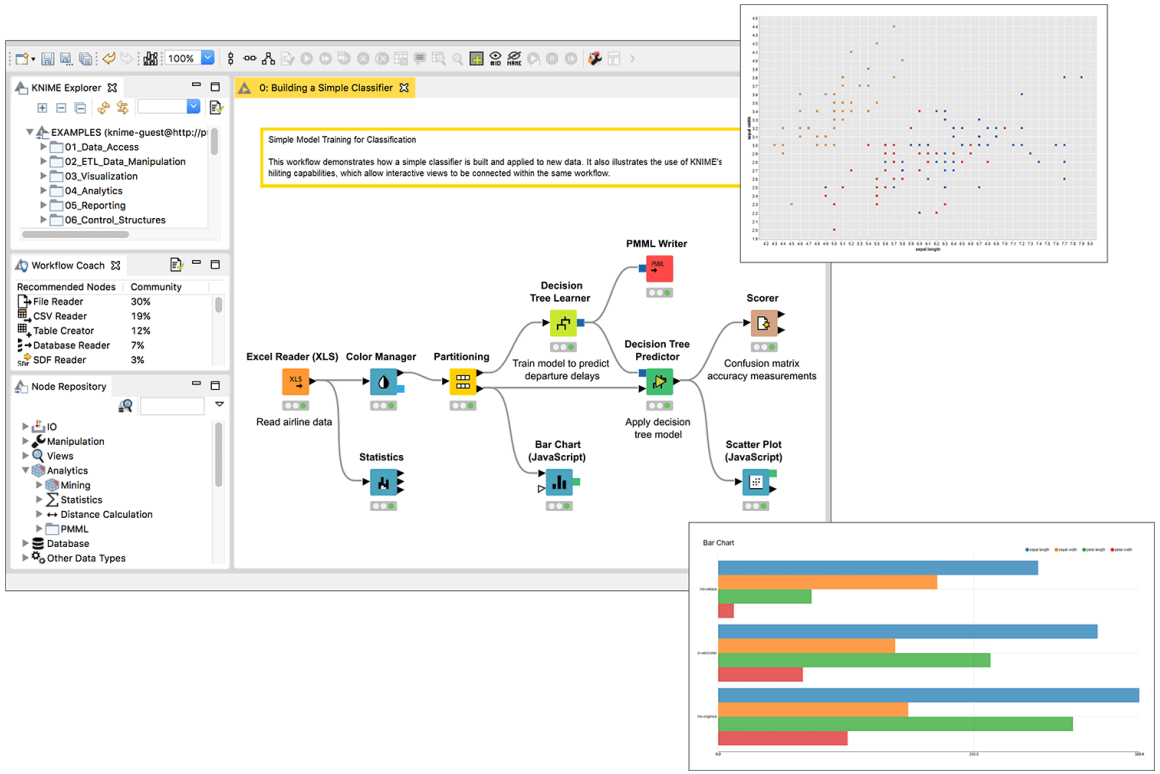
Open for Innovation ® KNIME

Brief Detour:
What is KNIME Analytics Platform?

© 2019 KNIME AG. All Rights Reserved.

What is KNIME Analytics Platform?

- A tool for data analysis, manipulation, visualization, and reporting
- Based on the graphical programming paradigm

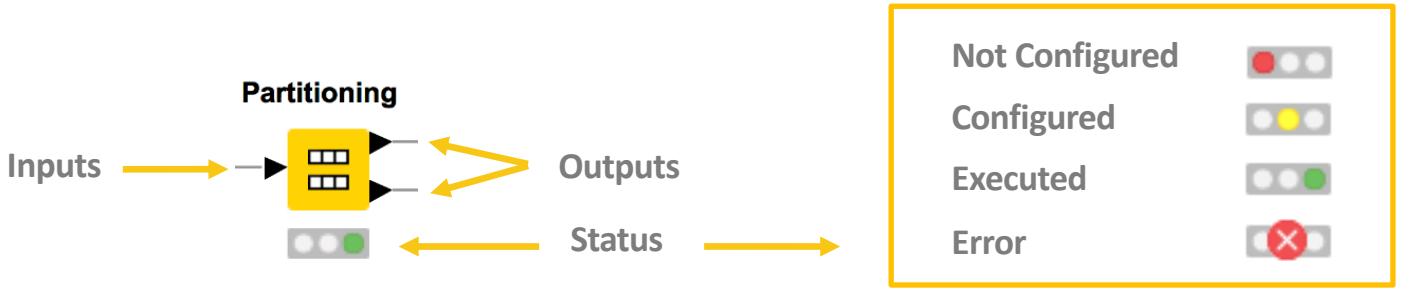


© 2019 KNIME AG. All Rights Reserved.

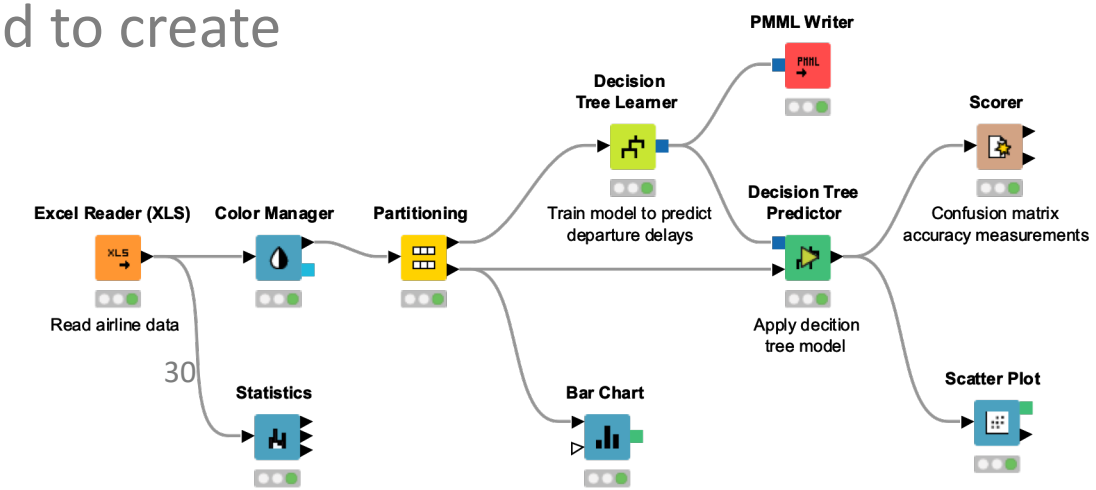
Open for Innovation ® KNIME

Visual KNIME Workflows

NODES perform tasks on data



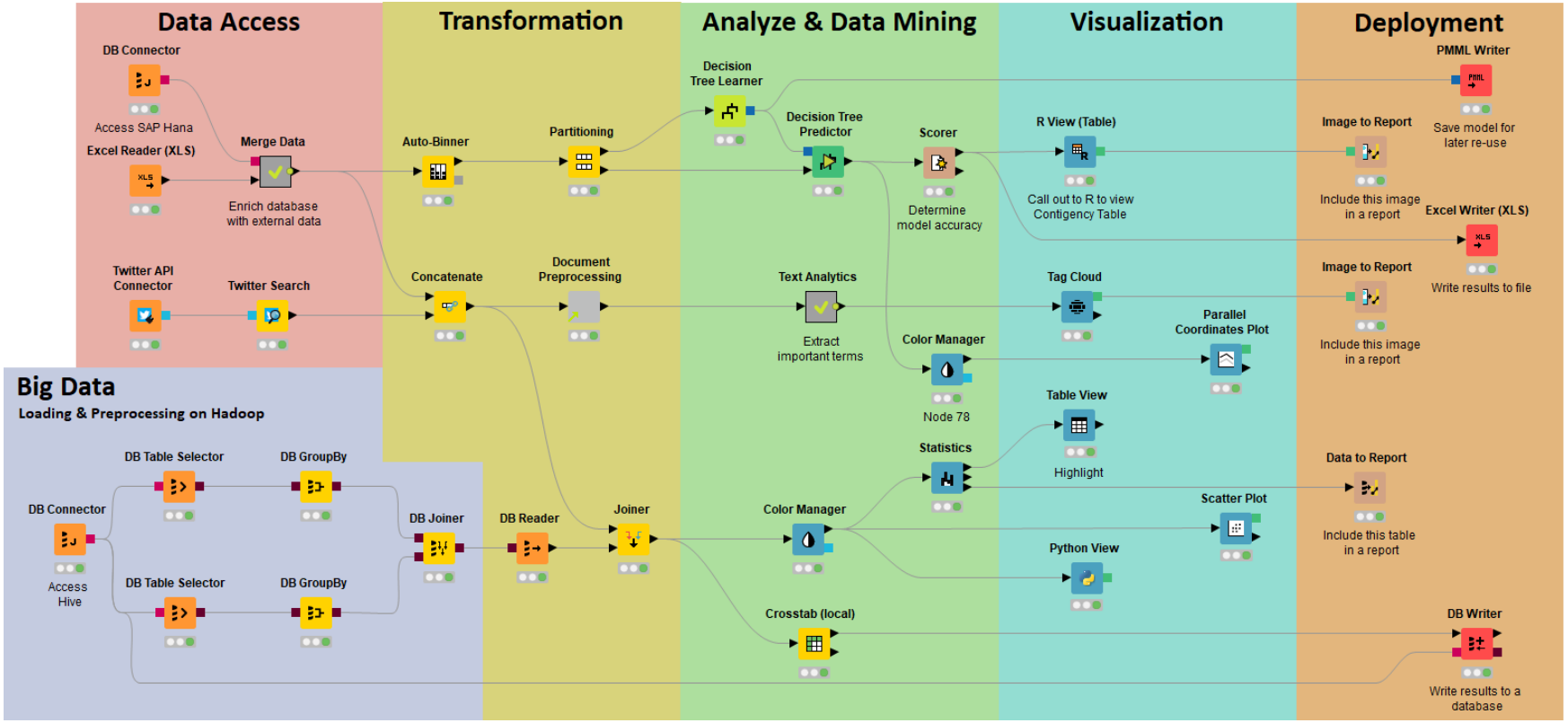
Nodes are combined to create **WORKFLOWS**



© 2019 KNIME AG. All Rights Reserved.

Open for Innovation KNIME

Over 2000 Native and Embedded Nodes Included:



Data Access
MySQL, Oracle, ...
SAS, SPSS, ...
Excel, Flat, ...
Hive, Impala, ...
XML, JSON, PMML
Text, Doc, Image, ...
Web Crawlers
Industry Specific
Community / 3rd

Transformation
Row
Column
Matrix
Text, Image
Time Series
Java
Python
Community / 3rd

Analysis & Mining
Statistics
Data Mining
Machine Learning
Web Analytics
Text Mining
Network Analysis
Social Media Analysis
R, Weka, Python
Community / 3rd

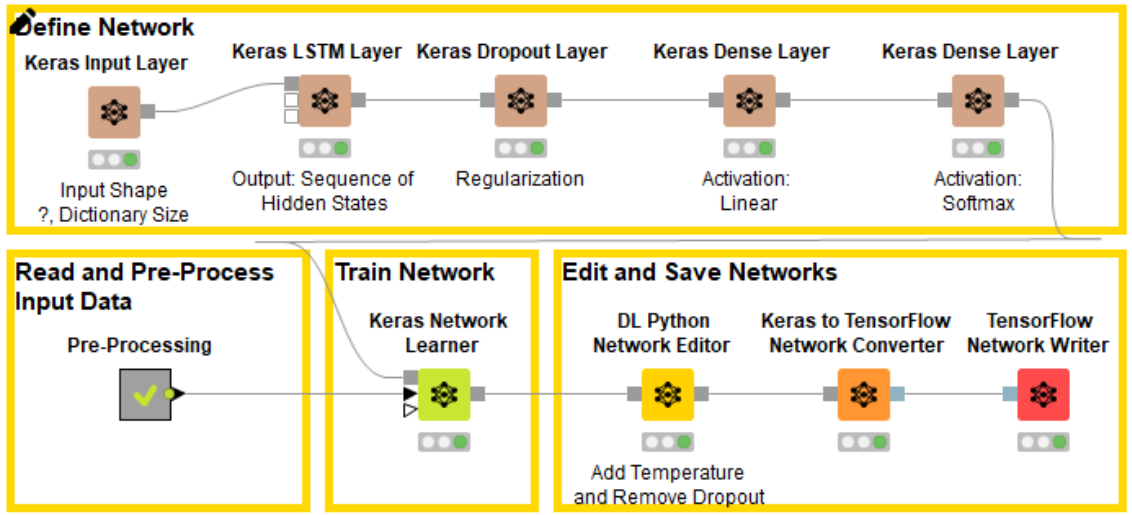
Visualization
R
JFreeChart
JavaScript
Plotly
Community / 3rd

Deployment
via BIRT
PMML
XML, JSON
Databases
Excel, Flat, etc.
Text, Doc, Image
Industry Specific
Community / 3rd

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ® KNIME

Model training workflow with KNIME Analytics Platform

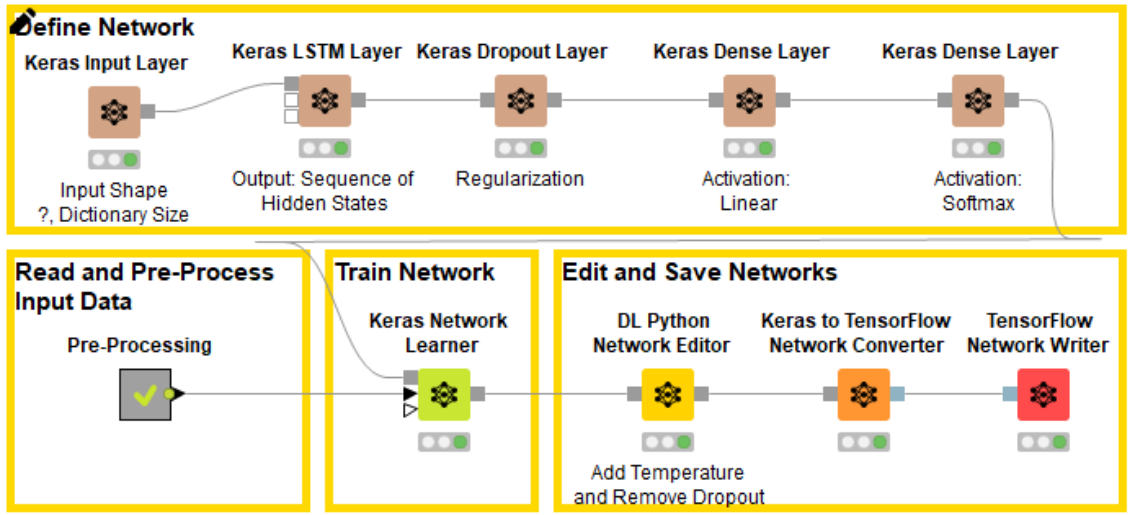


- Read in and transform our input dataset
- Define the structure of the neural network
- Train the model using Keras
- Do a bit of post-training model editing to introduce *temperature*
- Convert model to TensorFlow and save for use in deployment workflow

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation
KNIME

Neural network: Code-free



Layers of our neural network:

- Input layer of size [?, 95]
- LSTM for sequence analysis
- Dropout layer to prevent overfitting
- Dense layer 1 – linear activation
- Dense layer 2 (output) – Softmax activation

Once we've defined the network structure and pre-processed our input, we're ready to train.

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ® KNIME

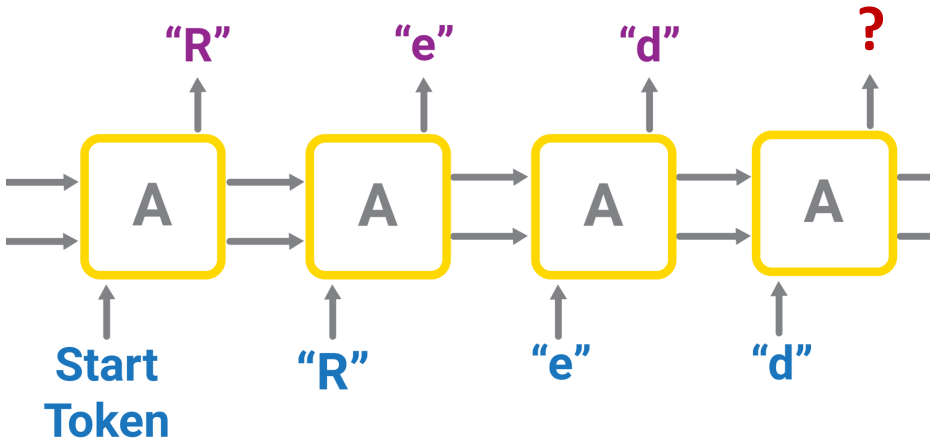
Demo:
Code-Free Neural Networks

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ®
KNIME

Now that we have a model, how can we use it?

- Output: Probability distribution over the different characters
  - Option 1: Use the character with the highest probability
    => Same start token leads to same mountain name
  - Option 2: Sample character according to the probability distribution
    => Different result for each execution

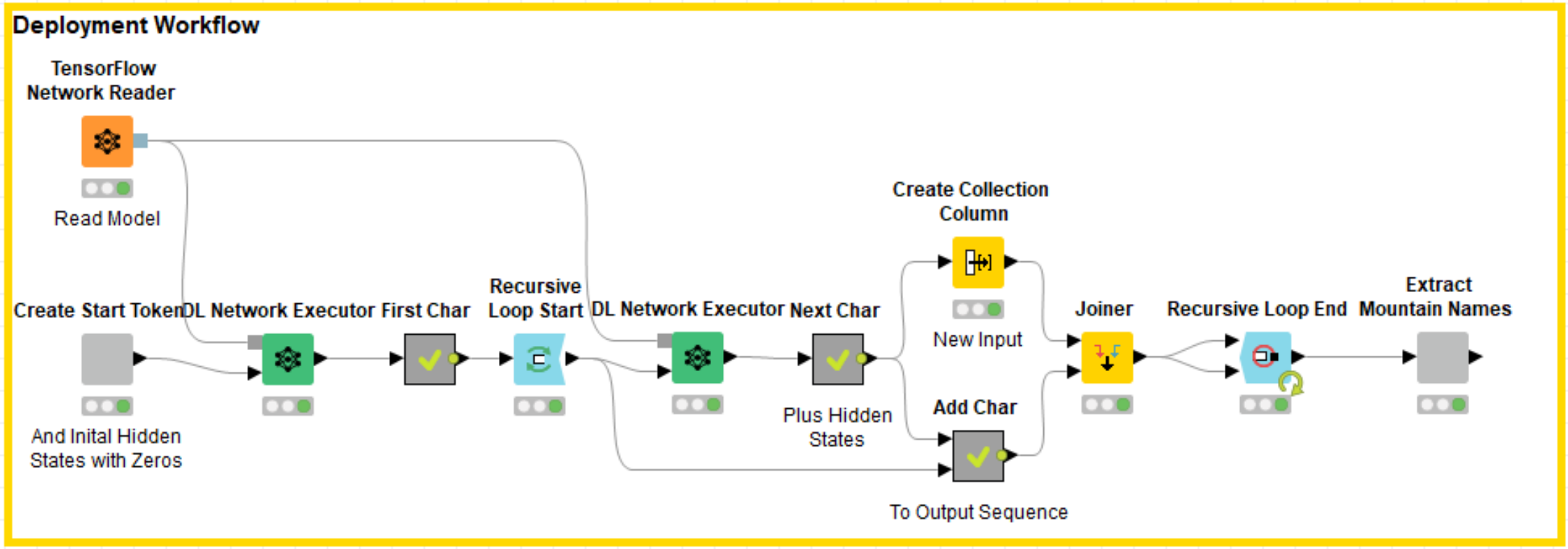


Character	Probability
<space>	0,2
x	0,002
m	0,105
z	0,0001
h	0,19

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation KNIME

Now that we have a model, how do we use it?



© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ® KNIME

Temperature parameter to control the creativity of the model

- Change the confidence of the model after training
- Lower temperatures $\tau < 1$:
  - Downplays the less likely predictions
- Higher temperatures $\tau > 1$:
  - Emphasis the least likely predictions
- Example: Probabilities (0.8, 0.1, 0.05, 0.05)
  - $\boldsymbol{\tau}$ = 1 => (0.8, 0.1, 0.05, 0.05)
  - $\boldsymbol{\tau}$ = 10 => (0.3, 0.24, 0.22, 0.22)
  - $\boldsymbol{\tau}$ = 0.1 => (0.99, 0.93e-11, 0.90e-14, 0.90e-14)

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ® KNIME

Temperature parameter to control the creativity of the model

- Change the confidence of the network after training
- Softmax

$$P(y = j|x) = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \text{ with } z_i = x^T w_i$$

- Softmax with temperature $\boldsymbol{\tau}$

$$P(y = j|x) = \frac{e^{z_j/\tau}}{\sum_{k=1}^{K} e^{z_k/\tau}} \text{ with } z_i = x^T w_i.$$

$\boldsymbol{\tau} < 1$ tends to more conservative

$\boldsymbol{\tau} > 1$ tends to more divers

© 2019 KNIME AG. All Rights Reserved.

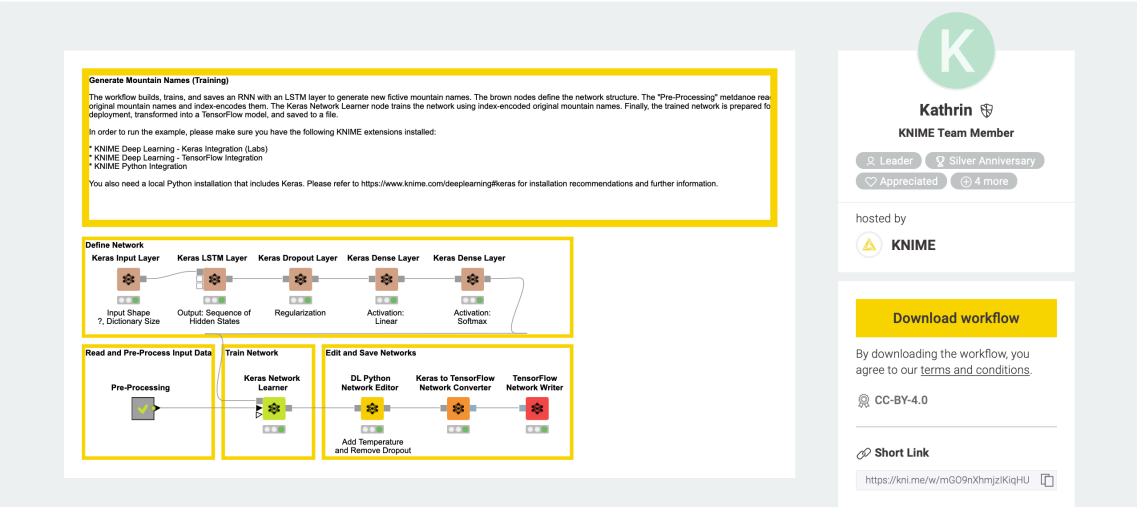
Open for Innovation ® KNIME

Results: Inspiration for the marketing team

Note that this doesn't have to be mountain names. It could have been any sort of text.

Try it out using your dataset:

Generate Mountain Names (Training)



https://kni.me/w/mGO9nXhmjzIKiqHU

	combined string
■	North Dalie Mountain
■	Glatfook Mountain
■	Mount Crapsland
■	Roblin Hill
■	Barlock Mountain
■	Frisby Mountain
■	Skodleger Peak
■	Casterland Mountain
■	Willow Knob
■	Wagaunovan Mountain
■	Little Ridge
■	Widdhern
■	Macomber Knya Hill
■	Roybel Butte
■	Cales Mountain
■	Welberly Hill

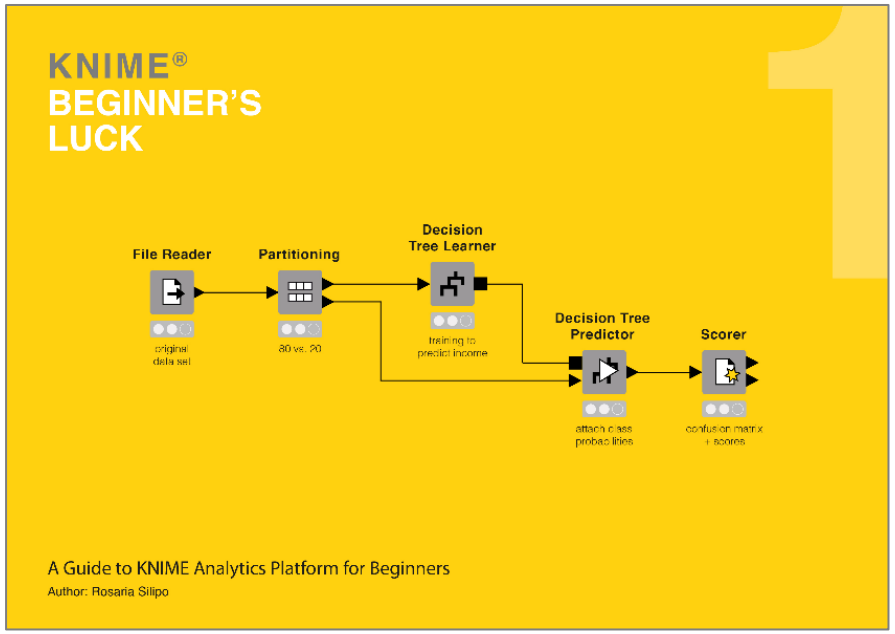
© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ® KNIME

KNIME Beginner's Luck

Book downloadable from **KNIME Press**
https://www.knime.com/knimepress

Get your free copy with the promotion code: **VILNIUS-2019**



© 2019 KNIME AG. All Rights Reserved.

Open for Innovation
KNIME

Thank you!

Questions?

The KNIME® trademark and logo and OPEN FOR INNOVATION® trademark are used by KNIME AG under license from KNIME GmbH, and are registered in the United States. KNIME® is also registered in Germany.

The KNIME® trademark and logo and OPEN FOR INNOVATION® trademark are used by KNIME AG under license from KNIME GmbH, and are registered in the United States. KNIME® is also registered in Germany.

© 2019 KNIME AG. All Rights Reserved.

Open for Innovation ®
KNIME