

Entendendo Comportamento de Compra em Dados Temporais através de Redes Neurais Recorrentes



Kevin Takano

<http://github.com/takanovcs>

takano@icomp.ufam.edu.br, kevin.takano@linux3.com

- ▶ Estudantes, Graduados?

- ▶ Estudantes, Graduados?
- ▶ Redes Neurais

- ▶ Estudantes, Graduados?
- ▶ Redes Neurais
- ▶ Redes Neurais Recorrentes. LSTM. GRU.

- ▶ Estudantes, Graduados?
- ▶ Redes Neurais
- ▶ Redes Neurais Recorrentes. LSTM. GRU.
- ▶ Embedding.

- ▶ Estudantes, Graduados?
- ▶ Redes Neurais
- ▶ Redes Neurais Recorrentes. LSTM. GRU.
- ▶ Embedding.
- ▶ Rede Convolutacional. Rede Convolutacional Dilatada.

- ▶ Estudantes, Graduados?
- ▶ Redes Neurais
- ▶ Redes Neurais Recorrentes. LSTM. GRU.
- ▶ Embedding.
- ▶ Rede Convolutacional. Rede Convolutacional Dilatada.
- ▶ Keras

- ▶ Estudantes, Graduados?
- ▶ Redes Neurais
- ▶ Redes Neurais Recorrentes. LSTM. GRU.
- ▶ Embedding.
- ▶ Rede Convolutacional. Rede Convolutacional Dilatada.
- ▶ Keras
- ▶ Curva ROC, AUC.

- ▶ Estudantes, Graduados?
- ▶ Redes Neurais
- ▶ Redes Neurais Recorrentes. LSTM. GRU.
- ▶ Embedding.
- ▶ Rede Convolutacional. Rede Convolutacional Dilatada.
- ▶ Keras
- ▶ Curva ROC, AUC.
- ▶ *Data Augmentation*

- ▶ Estudantes, Graduados?
- ▶ Redes Neurais
- ▶ Redes Neurais Recorrentes. LSTM. GRU.
- ▶ Embedding.
- ▶ Rede Convolutacional. Rede Convolutacional Dilatada.
- ▶ Keras
- ▶ Curva ROC, AUC.
- ▶ *Data Augmentation*
- ▶ *Oversampling e Undersampling.*

Como modelar dados temporais?

Redes Neurais Recorrentes

Geração de texto baseado em Alice No País Das Maravilhas

Prevendo comportamento de usuários através de RNN

Introdução

Datasets

Processamento de Dados

Arquitetura

Resultados

Como modelar dados temporais?

Redes Neurais Recorrentes

Geração de texto baseado em Alice No País Das Maravilhas

Prevendo comportamento de usuários através de RNN

Introdução

Datasets

Processamento de Dados

Arquitetura

Resultados

- Diversos campos de hoje possuem dados temporais: como **processamento de vídeo, de texto e de áudio**.

- ▶ Diversos campos de hoje possuem dados temporais: como **processamento de vídeo, de texto e de áudio**.
- ▶ As redes neurais recorrentes são uma das principais redes usadas para modelar dados temporais.

- ▶ Diversos campos de hoje possuem dados temporais: como **processamento de vídeo, de texto e de áudio**.
- ▶ As redes neurais recorrentes são uma das principais redes usadas para modelar dados temporais.
- ▶ RNNs tem uma **dimensão a mais** que possibilita a entrada de um dado que é **dependente** do anterior. (Dado temporal)

- ▶ Diversos campos de hoje possuem dados temporais: como **processamento de vídeo, de texto e de áudio**.
- ▶ As redes neurais recorrentes são uma das principais redes usadas para modelar dados temporais.
- ▶ RNNs tem uma **dimensão a mais** que possibilita a entrada de um dado que é **dependente** do anterior. (Dado temporal)
- ▶ OBS: Atualmente os modelos de RNN **não** são estado da arte nos campos mencionados mas ainda são usados em outras áreas. Ex: modelo com attention.

Como modelar dados temporais?

Redes Neurais Recorrentes

Geração de texto baseado em Alice No País Das Maravilhas

Prevendo comportamento de usuários através de RNN

Introdução

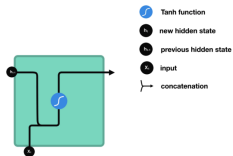
Datasets

Processamento de Dados

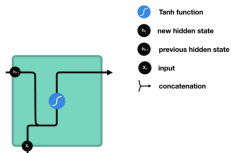
Arquitetura

Resultados

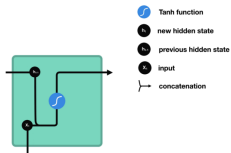
► Célula básica



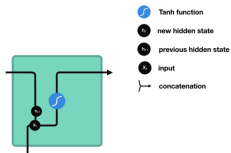
► Célula básica



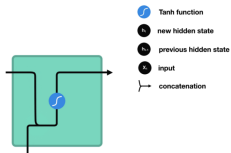
► Célula básica



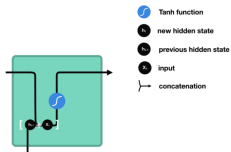
► Célula básica



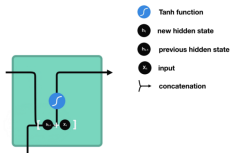
► Célula básica



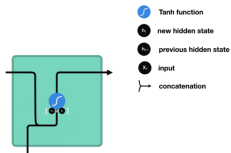
► Célula básica



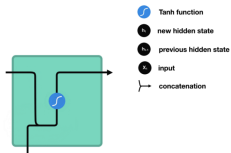
► Célula básica



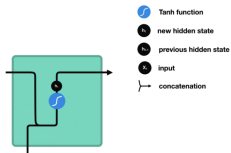
► Célula básica



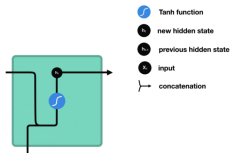
► Célula básica



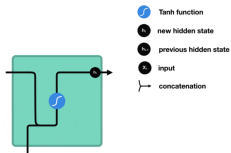
► Célula básica



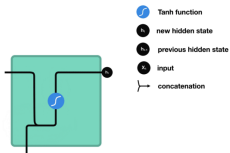
► Célula básica



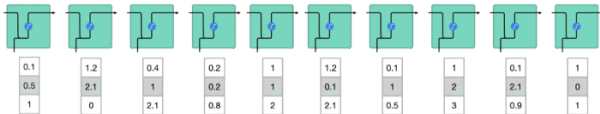
► Célula básica



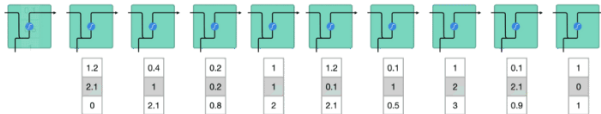
- Célula básica



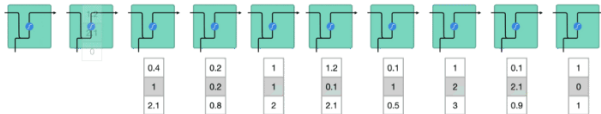
► RNN: Sequência de Células



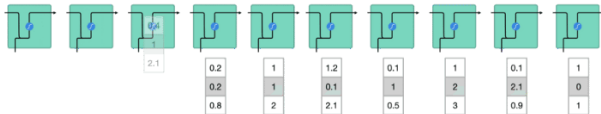
► RNN: Sequência de Células



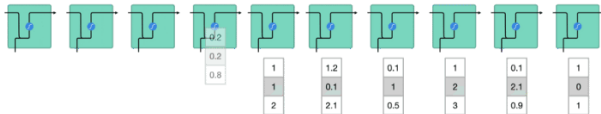
► RNN: Sequência de Células



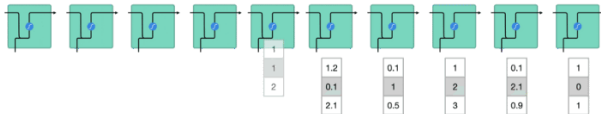
► RNN: Sequência de Células



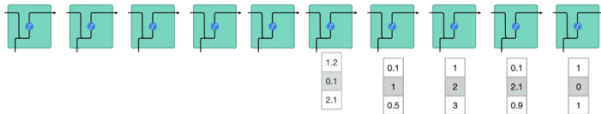
► RNN: Sequência de Células



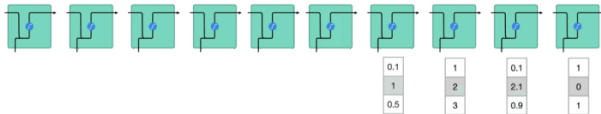
► RNN: Sequência de Células



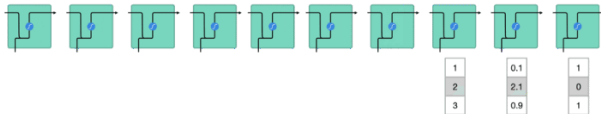
► RNN: Sequência de Células



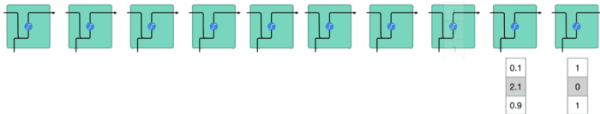
► RNN: Sequência de Células



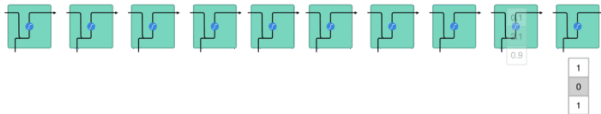
► RNN: Sequência de Células



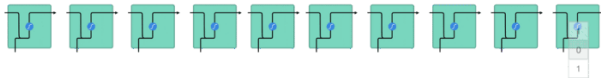
► RNN: Sequência de Células



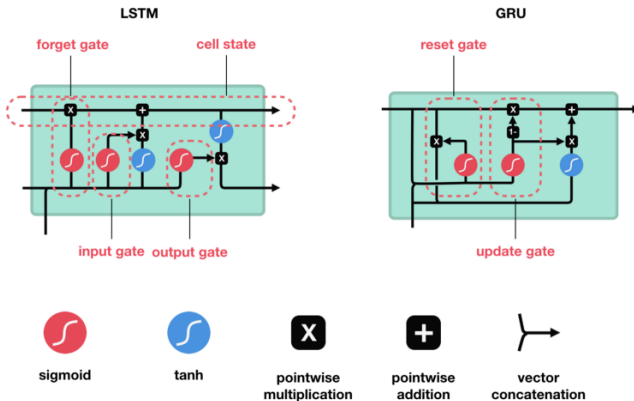
► RNN: Sequência de Células



► RNN: Sequência de Células



- Vanish Gradient Problem. LSTM & GRU.



Como modelar dados temporais?

Redes Neurais Recorrentes

Geração de texto baseado em Alice No País Das Maravilhas

Prevendo comportamento de usuários através de RNN

Introdução

Datasets

Processamento de Dados

Arquitetura

Resultados

ALICE'S ADVENTURES IN WONDERLAND

Lewis Carroll

THE MILLENNIUM FULCRUM EDITION 3.0

CHAPTER I.

Down the Rabbit-Hole Alice was beginning to get very tired of sitting by her sister on the bank, and of having nothing to do: once or twice she had peeped into the book her sister was reading, but it had no pictures or conversations init, 'and what is the use of a book,' thought Alice 'without pictures orconversations?' So she was considering in her own mind (as well as she could, for the hot day made her feel very sleepy and stupid), whether the pleasure of making a daisy-chain would be worth the trouble of getting up and picking the daisies, when suddenly a White Rabbit with pink eyes ran close by her.

ALICE'S ADVENTURES IN WONDERLAND

Lewis Carroll

THE MILLENNIUM FULCRUM EDITION 3.0

CHAPTER I.

Down the Rabbit-Hole Alice was beginning to get very tired of sitting by her sister on the bank, and of having nothing to do: once or twice she had peeped into the book her sister was reading, but it had no pictures or conversations init, 'and what is the use of a book,' thought Alice 'without pictures orconversations?' So she was considering in her own mind (as well as she could, for the hot day made her feel very sleepy and stupid), whether the pleasure of making a daisy-chain would be worth the trouble of getting up and picking the daisies, when suddenly a White Rabbit with pink eyes ran close by her.

ALICE'S ADVENTURES IN WONDERLAND

Lewis Carroll

THE MILLENNIUM FULCRUM EDITION 3.0

CHAPTER I.

input

output

Down the Rabbit-Hole Alice was beginning to get very tired of sitting by her sister on the bank, and of having nothing to do: once or twice she had peeped into the book her sister was reading, but it had no pictures or conversations init, 'and what is the use of a book,' thought Alice 'without pictures orconversations?' So she was considering in her own mind (as well as she could, for the hot day made her feel very sleepy and stupid), whether the pleasure of making a daisy-chain would be worth the trouble of getting up and picking the daisies, when suddenly a White Rabbit with pink eyes ran close by her.

ALICE'S ADVENTURES IN WONDERLAND

Lewis Carroll

THE MILLENNIUM FULCRUM EDITION 3.0

CHAPTER I.

input

output

Down the Rabbit-Hole Alice was beginning to get very tired of sitting by her sister on the bank, and of having nothing to do: once or twice she had peeped into the book her sister was reading, but it had no pictures or conversations in it, 'and what is the use of a book,' thought Alice 'without pictures or conversations?' So she was considering in her own mind (as well as she could, for the hot day made her feel very sleepy and stupid), whether the pleasure of making a daisy-chain would be worth the trouble of getting up and picking the daisies, when suddenly a White Rabbit with pink eyes ran close by her.

ALICE'S ADVENTURES IN WONDERLAND

Lewis Carroll

THE MILLENNIUM FULCRUM EDITION 3.0

CHAPTER I.

input

output

Down the Rabbit-Hole Alice was beginning to get very tired of sitting by her sister on the bank, and of having nothing to do: once or twice she had peeped into the book her sister was reading, but it had no pictures or conversations init, 'and what is the use of a book,' thought Alice 'without pictures orconversations?' So she was considering in her own mind (as well as she could, for the hot day made her feel very sleepy and stupid), whether the pleasure of making a daisy-chain would be worth the trouble of getting up and picking the daisies, when suddenly a White Rabbit with pink eyes ran close by her.

ALICE'S ADVENTURES IN WONDERLAND

Lewis Carroll

THE MILLENNIUM FULCRUM EDITION 3.0

CHAPTER I.

input

output

Down the Rabbit-Hole Alice was beginning to get very tired of sitting by her sister on the bank, and of having nothing to do: once or twice she had peeped into the book her sister was reading, but it had no pictures or conversations init, 'and what is the use of a book,' thought Alice 'without pictures orconversations?' So she was considering in her own mind (as well as she could, for the hot day made her feel very sleepy and stupid), whether the pleasure of making a daisy-chain would be worth the trouble of getting up and picking the daisies, when suddenly a White Rabbit with pink eyes ran close by her.

ALICE'S ADVENTURES IN WONDERLAND

Lewis Carroll

THE MILLENNIUM FULCRUM EDITION 3.0

CHAPTER I.

input

output

Down the Rabbit-Hole Alice was beginning to get very tired of sitting by her sister on the bank, and of having nothing to do: once or twice she had peeped into the book her sister was reading, but it had no pictures or conversations init, 'and what is the use of a book,' thought Alice 'without pictures orconversations?'So she was considering in her own mind (as well as she could, for the hot day made her feel very sleepy and stupid), whether the pleasure of making a daisy-chain would be worth the trouble of getting up and picking the daisies, when suddenly a White Rabbit with pink eyes ran close by her.

► Arquitetura da Rede:

```
model = Sequential()  
model.add(GRU(64, input_shape=(X.shape[1], X.shape[2])))  
model.add(Dropout(0.2))  
model.add(Dense(y.shape[1], activation='softmax'))  
model.compile(loss='categorical_crossentropy',  
optimizer='adam')
```

- ▶ Exemplo de entrada: said, tossing his head contemptuously. 'i dare say you never even spoke to time!' 'perhaps n

- ▶ Exemplo de entrada: said, tossing his head contemptuously. 'i dare say you never even spoke to time!' 'perhaps n
- ▶ Exemplo de saída: hatter said, tossing his head contemptuously.'i dare say you never even spoke to time!'perhaps n o to the toue ' said the cocp.'ih the woeee to the cotee to the toeee an c den toet toeee to the toeee to the toue lo tou do the soue ' said the cocp.''to the toeee to thet the soeee to the toeee to tee touee the souee io tou dooe the sone toee the sar the soeee to the soeee "

Como modelar dados temporais?

Redes Neurais Recorrentes

Geração de texto baseado em Alice No País Das Maravilhas

Prevendo comportamento de usuários através de RNN

Introdução

Datasets

Processamento de Dados

Arquitetura

Resultados

Como modelar dados temporais?

Redes Neurais Recorrentes

Geração de texto baseado em Alice No País Das Maravilhas

Prevendo comportamento de usuários através de RNN

Introdução

Datasets

Processamento de Dados

Arquitetura

Resultados

- ▶ Hoje sabe-se que boa parte das lojas de varejos online guardam dados de usuário.

- ▶ Hoje sabe-se que boa parte das lojas de varejos online guardam dados de usuário.
- ▶ Por questões de marketing, sabe-se que é muito vantajoso entender o interesse dos consumidores.
- ▶ Aumento de 2 a 11%

- ▶ Pergunta de pesquisa:

- ▶ Pergunta de pesquisa:
- ▶ *Considerando que em suas bases as lojas de e-commerce guardam os dados de forma histórica:*
 - ▶ Podemos prever qual será seu próximo evento ?
 - ▶ Podemos prever se ele irá ser um comprador ou não?

- Formato de dados de usuários em sequência de eventos (sessão de cliques):

v v v v t a t t t v v v v v v v a t

v: view, t: transaction, a: add-to-cart

- Formato de dados de usuários em sequência de eventos (sessão de cliques):

v v v v t a t t t v v v v v v v a t

v: view, t: transaction, a: add-to-cart

- Uma RNN pode encontrar um padrão nessa sequência para prever o próximo evento.

- ▶ Formato de dados de usuários em sequência de eventos (sessão de cliques):

v v v v t a t t t v v v v v v v a t

v: view, t: transaction, a: add-to-cart

- ▶ Uma RNN pode encontrar um padrão nessa sequência para prever o próximo evento.
- ▶ Através disso poderemos saber se usuário irá comprar ou não.

- ▶ Nesse trabalho, considere apenas a classificação de compradores e não-compradores. Isto é:
 - ▶ Quem apenas gerou evento de *view* e *add-to-cart* é não comprador.
 - ▶ Quem gerou evento de *transaction* é comprador.

Como modelar dados temporais?

Redes Neurais Recorrentes

Geração de texto baseado em Alice No País Das Maravilhas

Prevendo comportamento de usuários através de RNN

Introdução

Datasets

Processamento de Dados

Arquitetura

Resultados

- ▶ Dois datasets usados: Retail Rocket, RecSys 2015.

- ▶ Dois datasets usados: Retail Rocket, RecSys 2015.
- ▶ O conjunto de dados do RecSys 2015 consiste em 9,2 milhões de sessões de cliques. Retail Rocket contém 1.4 milhões de sessões.

- ▶ Dois datasets usados: Retail Rocket, RecSys 2015.
- ▶ O conjunto de dados do RecSys 2015 consiste em 9,2 milhões de sessões de cliques. Retail Rocket contém 1.4 milhões de sessões.
- ▶ RecSys contém apenas 5% e o Retail Rocket tem apenas 0.7% de sessões de compras. (*Desbalanceado*)

- ▶ Dois datasets usados: Retail Rocket, RecSys 2015.
- ▶ O conjunto de dados do RecSys 2015 consiste em 9,2 milhões de sessões de cliques. Retail Rocket contém 1.4 milhões de sessões.
- ▶ RecSys contém apenas 5% e o Retail Rocket tem apenas 0.7% de sessões de compras. (*Desbalanceado*)
- ▶ Isto torna o problema muito fácil de confundir o modelo.

- ▶ Dois datasets usados: Retail Rocket, RecSys 2015.
- ▶ O conjunto de dados do RecSys 2015 consiste em 9,2 milhões de sessões de cliques. Retail Rocket contém 1.4 milhões de sessões.
- ▶ RecSys contém apenas 5% e o Retail Rocket tem apenas 0.7% de sessões de compras. (*Desbalanceado*)
- ▶ Isto torna o problema muito fácil de confundir o modelo.
- ▶ RecSys dataset não contém eventos *add-to-cart*.

Como modelar dados temporais?

Redes Neurais Recorrentes

Geração de texto baseado em Alice No País Das Maravilhas

Prevendo comportamento de usuários através de RNN

Introdução

Datasets

Processamento de Dados

Arquitetura

Resultados

Técnicas e Processamento de Dados:

- ▶ O tamanho das sessões em e-commerce variam de e 1% (cerca de 90% das sessões) até 11000. (RecSys).

Técnicas e Processamento de Dados:

- ▶ O tamanho das sessões em e-commerce variam de e 1% (cerca de 90% das sessões) até 11000. (RecSys).
- ▶ T1: Pegar uma janela dos últimos n timesteps. Tamanho 100.

Técnicas e Processamento de Dados:

- ▶ O tamanho das sessões em e-commerce variam de e 1% (cerca de 90% das sessões) até 11000. (RecSys).
- ▶ T1: Pegar uma janela dos últimos n timesteps. Tamanho 100.
- ▶ T2: Criação de mais dados de entrada através de *data augmentation*

Técnicas e Processamento de Dados:

- ▶ O tamanho das sessões em e-commerce variam de e 1% (cerca de 90% das sessões) até 11000. (RecSys).
- ▶ T1: Pegar uma janela dos últimos n timesteps. Tamanho 100.
- ▶ T2: Criação de mais dados de entrada através de *data augmentation*
- ▶ T3: Usar *undersampling* e *oversampling* para resolver o problema de balanceamento.

Técnicas e Processamento de Dados:

- ▶ Formato de dados original:
 - ▶ Evento | ItemId | CategoriaID | Timestamp | Custo

Técnicas e Processamento de Dados:

- ▶ Formato de dados original:
 - ▶ Evento | ItemId | CategoriaID | Timestamp | Custo
- ▶ Processado:
 - ▶ Evento | ItemId | CategoriaID | HoraDaCompra |
DiaDaSemana | Mes | SemanaDoAno |
TempoEntreCliques(*dwelltime*) | Custo

Técnicas e Processamento de Dados:

- ▶ Formato de dados original:
 - ▶ Evento | ItemId | CategoriaID | Timestamp | Custo
- ▶ Processado:
 - ▶ Evento | ItemId | CategoriaID | HoraDaCompra | DiaDaSemana | Mes | SemanaDoAno | TempoEntreCliques(*dwelltime*) | Custo
- ▶ Com exceção de *dwelltime* e o *custo* todos os campos receberam um Embedding na arquitetura final.

Técnicas e Processamento de Dados:

- ▶ Formato de dados original:
 - ▶ Evento | ItemId | CategoriaID | Timestamp | Custo
- ▶ Processado:
 - ▶ Evento | ItemId | CategoriaID | HoraDaCompra |
DiaDaSemana | Mes | SemanaDoAno |
TempoEntreCliques(*dwelltime*) | Custo
- ▶ Com exceção de *dwelltime* e o *custo* todos os campos receberam um Embedding na arquitetura final.
- ▶ Custo e *dwelltime* são único dois campos não categóricos, por isso foram *normalizados*.

Como modelar dados temporais?

Redes Neurais Recorrentes

Geração de texto baseado em Alice No País Das Maravilhas

Prevendo comportamento de usuários através de RNN

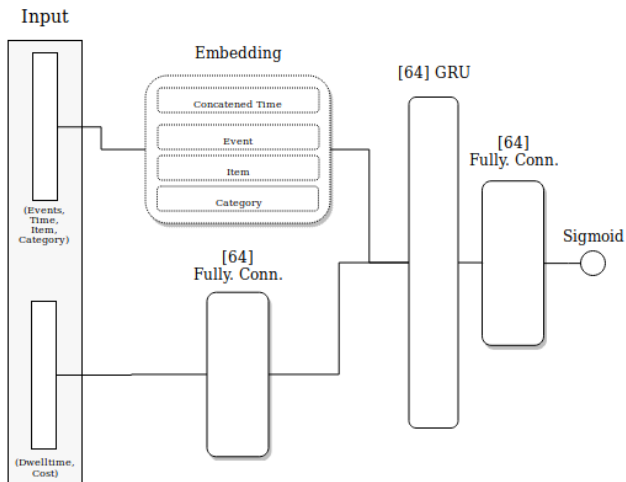
Introdução

Datasets

Processamento de Dados

Arquitetura

Resultados



► Dilated Convolution.

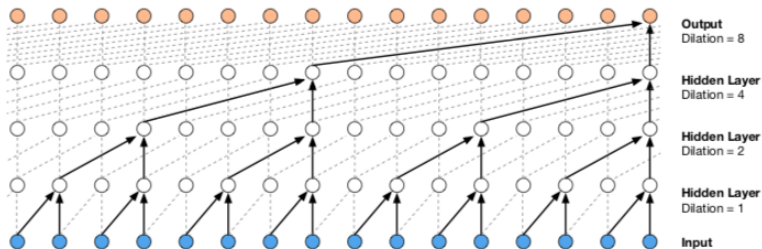
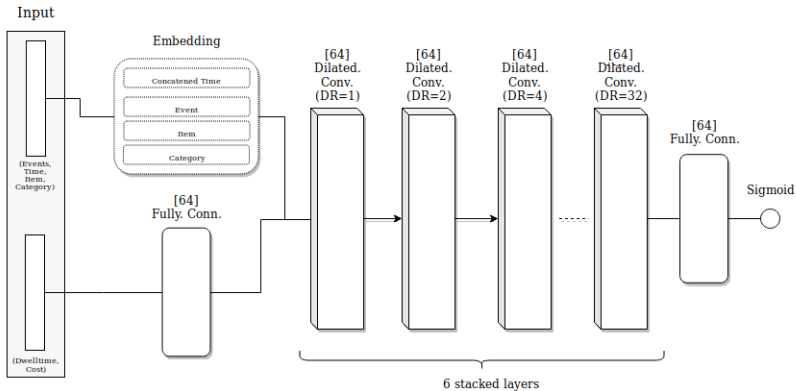


Figure 3: Visualization of a stack of *dilated* causal convolutional layers.



Como modelar dados temporais?

Redes Neurais Recorrentes

Geração de texto baseado em Alice No País Das Maravilhas

Prevendo comportamento de usuários através de RNN

Introdução

Datasets

Processamento de Dados

Arquitetura

Resultados

Resultados a priori (RecSys):

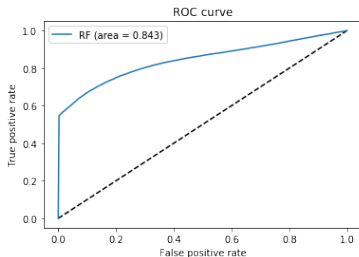


Figure: GRU

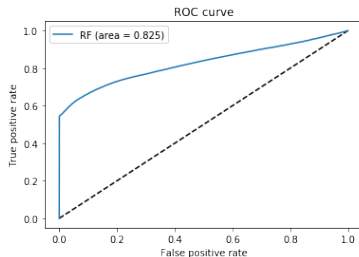


Figure: Dilated Convolution

Resultados a priori (RetailRocket):

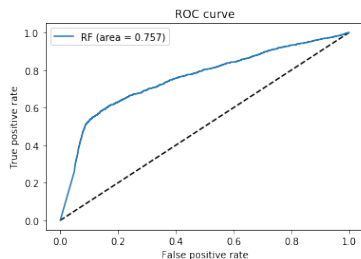


Figure: GRU

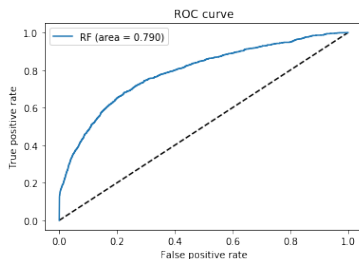


Figure: Dilated Convolution

Valeuzão!