

Long/Short Term Memory (LSTM)

Predicting blood's glucose level using LSTM network

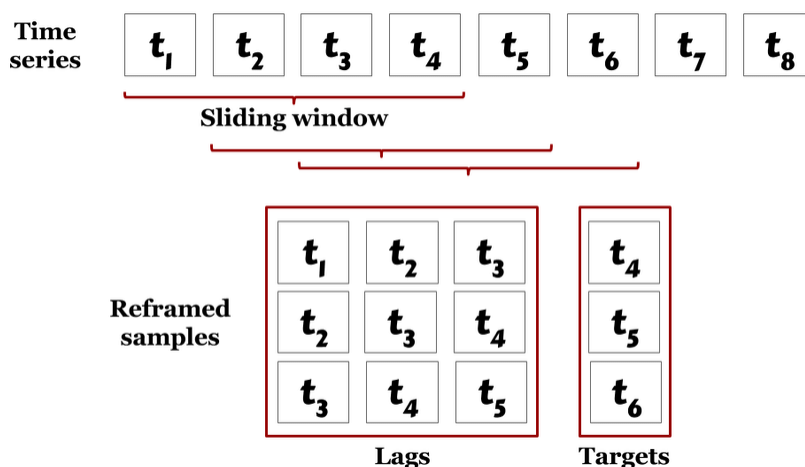
امیر محمد فلاح (۹۷۰۱۲۲۶۸۰۰۱۴)

دیتاست:

کلینیک ثبت اطلاعات دیابت نوع یک (The T1D Exchange) ، به جمع آوری اطلاعات ۲۰۳ بیمار میانسال مبتلا به این نوع دیابت پرداخته است. به عنوان بخشی از داده های ثبت شده در ثبتنام، اطلاعات مرتبط با رویداد هیپوگلیسمی نیز جمع آوری شده است. اطلاعات مورد استفاده از دیتاست در این پروژه ستون glucose بود که نشان دهنده رکورد های میزان قند خون هر بیمار بوده که هر ۵ دقیقه جمع آوری شده است.

PtID	Glucose	SHFacRecSkipMeal	SHFacRecAteLess	SHFacRecMiscalcCarb	SHFacRecInsAhead	SHFacRecExercise	SHFacRecInsAdminError
199	200	0	0	1	0	1	0
199	195	0	0	1	0	1	0
199	236	0	0	1	0	1	0
199	210	0	0	1	0	1	0
199	196	0	0	1	0	1	0
199	230	0	0	1	0	1	0
199	239	0	0	1	0	1	0
199	250	0	0	1	0	1	0

در این پروژه، هدف ما ارائه روشی است که بتوانیم از شبکه های عصبی حافظه طولانی کوتاه مدت و پیشخور عمیق به منظور پیشبینی میزان قند خون استفاده کنیم. به منظور تحقق پیشبینی دقیق، مدل پیشنهادی بدین صورت است که ابتدا مجموعه داده سری زمانی به صورت داده تحت یادگیری با نظارت تبدیل میشود. به این صورت که داده های سری زمانی به صورت نگاشتی از وضعیت حال حاضر به داده های آینده داشته باشند. برای این کار از شیوه Sliding Window استفاده شده است .



روش sliding window برای تبدیل داده های سری زمانی به داده های supervised استفاده می شود. نحوه عملکرد آن به این صورت است که به تعداد عدد ثابت lag observation، تعداد قدم های قبلی را به داده ی ورودی اضافه می کند. و خروجی را قدم بعدی قرار می دهد. بر این اساس می توان نگاشتی از X ها به y ها بر روی دیتاست پیدا کرد.

مدل شبکه عصبی پیشنهادی:

شبکه عصبی بازگشتی روشی است که برای مدل کردن داده های متوالی استفاده شده است. عموماً، شبکه عصبی بازگشتی از شبکه پیشفرض توسط اتصال خروجی عصبی به ورودی هایش به دست آمده است. این شبکه ها دارای یک نوع حافظه هستند که اطلاعاتی تاکنون دیده اند را ضبط میکنند. در تئوری اینطور به نظر میرسد که این شبکه ها میتوانند اطلاعات موجود در یک دنباله طولانی را ضبط و از آنها استفاده کنند اما در عمل اینطور نیست و محدود هستند، به این صورت که فقط اطلاعات چند گام قبل را ضبط میکنند. برخلاف شبکه های معمولی که از پارامترهای متفاوتی در هر لایه استفاده میکنند، یک شبکه عصبی بازگشتی پارامترهای مشابهی را بین همه گام های زمانی به اشتراک می گذارد. این بدین معنی است که در هر گام زمانی عملیات مشابهی را انجام میشود فقط ورودی ها متفاوت هستند. با این تکنیک تعداد کلی پارامترهایی که شبکه باید یاد بگیرد به شدت کاهش پیدا میکند. اصلیتین ویژگی این شبکه، حالت پنهان آن است که اطلاعاتی یک توالی را ذخیره میکند. شبکه های عصبی بازگشتی به این علت بازگشتی نامیده می شوند که خروجی هر لایه به محاسبات لایه های ماقبل آن وابسته است. به همین دلیل این شبکه ها دارای حافظه هستند که اطلاعات مربوط به داده های دیده شده را ذخیره میکنند. این شبکه ها درواقع کپی های متعددی از شبکه های عصبی معمولی هستند که کنار هم قرارگرفته اند و هرکدام پیغامی را به دیگری انتقال میدهند. توسعه معروف آن، استفاده از سلول حافظه طولانی کوتاه مدت (Long short term memory) است که بهتر وابستگی های بلندمدت را پیدا میکند. حافظه طولانی کوتاه مدت برای حل مشکل پدیده محوشدگی گرادیان در شبکه های عصبی بازگشتی به وجود آمدند که تغییر اصلی آن جایگزین کردن لایه میانی شبکه های عصبی بازگشتی با یک سلول LSTM است. بزرگترین ویژگی حافظه طولانی کوتاه مدت امکان یادگیری وابستگی بلندمدت است که توسط شبکه های عصبی بازگشتی امکانپذیر نبود. برای پیشبینی گام زمانی بعدی نیاز است که مقادیر وزن ها در شبکه به روزرسانی شوند که این کار مستلزم حفظ اطلاعات گامهای زمانی ابتدایی است. یک شبکه عصبی بازگشتی فقط میتواند تعداد محدودی از وابستگی های کوتاه مدت را یاد بگیرد، پس سری های زمانی بلندمدت قابل یادگیری توسط شبکه های عصبی بازگشتی نیستند و حافظه های طولانی کوتاه مدت میتوانند این وابستگی های بلندمدت را به درستی یاد بگیرند. در این شبکه سه گیت وجود دارد که از طریق آن شبکه نسبت به کنترل جریان داده درون خود اقدام میکند. این سه گیت عبارتند از گیت فراموشی، گیت به روز رسانی که به گیت ورودی هم معروف است و گیت خروجی. همچنین یک سلول حافظه نیز وجود دارد. علاوه بر این چهار مفهوم، دارای یک ورودی از حافظه پنهان و ورودی نیز بهره برده و خروجی تولید میکند که به دو بخش تقسیم میشود بخشی به گام زمانی بعد منتقل شده و بخشی نیز در صورت نیاز به تولید خروجی در گام زمانی فعلی مورد استفاده قرار میگیرد. مدل پیشنهادی ترکیبی از سلول LSTM و شبکه ی Fully Connected است که به صورت زیر می باشد.

Model: "sequential_1"		
Layer (type)	Output Shape	Param #
=====		
lstm_1 (LSTM)	(None, 60)	15360
dense_1 (Dense)	(None, 70)	4270
dense_2 (Dense)	(None, 80)	5680
dense_3 (Dense)	(None, 65)	5265
dense_4 (Dense)	(None, 15)	990
dense_5 (Dense)	(None, 1)	16
=====		
Total params: 31,581		
Trainable params: 31,581		
Non-trainable params: 0		

تمامی توابع فعال سازی استفاده شده در طول شبکه Relu بوده که فرمول آن به شرح زیر می باشد:

$$f(x) = x^+ = \max(0, x)$$

کانفیگ شبکه عصبی پیشنهادی:

- اپتیماایزر آدام با نرخ یادگیری ۰.۰۰۰۰۱
- Loss function: mean_absolute_error
- Epochs: 100