Validaide:

Temperature solution for pharmaceutical products transportation

James Wen, Wenxu Liu, Yuming Shi

{james_wen, wenxuliu, yumingshi} @ hsph.harvard.edu

May 09, 2022

Executive Summary

This report serves to provide temperature prediction profiles to facilitate optimal packaging decision-making for Validaide (https://www.validaide.com/). Predicting ambient temperature at airport tarmacs is essential, as cargos are especially susceptible to uncontrolled ambient temperature outside. Knowing the temperature also helps determine the packaging solution for different types of pharmaceutical products. Using an autoregressive model, and Long Short Term Memory (LSTM) neural networks, we created hourly - monthly ambient temperature predictions throughout 2020 for 15 representative international airports, using historical data from 2009 to 2019. The LSTM model performed better when the cyclical trend of temperatures were relatively steady across years, and was eventually applied to four airports: JNB, HYD, FRA, and DXB. Other 11 airports ended up employing the autoregressive model for ambient temperature predictions. The logistics and methodologies are provided by students in the project group 'Validaide' of MIT SCM.c51 course: Machine Learning Applications for Supply Chain Management.

INTRODUCTION

Shipping pharmaceuticals requires a high level of expertise to transport the products safely. In many cases, pharmaceutical products are sensitive and must be delivered in a timely manner. Temperature control is essential for maintaining the quality of the drugs. Some of the pharmaceutical products require as low as -80°C during transportation, and a slight change in the temperature, even of just 2 degrees, can damage the product. Products that do not maintain the right temperature could become ineffective, or even harmful for patients who receive them. Temperature controlled shipping is applied for pharmaceutical logistics as a part of the cold chain logistics, which refers to shipping, storing and distributing refrigerated and frozen goods[1]. Ensuring the drugs are well temperature-controlled throughout the transportation route is critical for pharmaceutical logistics.

The medicines and pharma products are usually transported in refrigerated or insulated containers. Some of the products that are viable in higher temperatures, such as 2°C - 8°C, can be kept in thermal blankets or softboxes, and stored in the cold storage units in airports waiting to be onboard. Others that need a much lower temperature to hold, such as the Pfizer vaccine, which requires a much lower degree at -70°C, require active or passive cooling. Active and passive cooling are containers that help keep products cool either electronically or with dry ice. If power runs out, the batteries built in the container will keep the cargo refrigerated, or the dry ice will passively keep it cool. Whether to use softboxes/thermal blankets, or active/passive cooling techniques, are determined by both the product type and the ambient temperature the product will pass by on the global transportation route.

A global transportation route of a pharmaceutical product usually involves the following: Warehouse storage, ground/road transportation through refrigerated trucks (also referred as reefer trucks) , ground handling at airport, and air transportation. In this process, the pharmaceutical product is most vulnerable during transfer handling at each airport. Specifically, the cargo is sensitive to ambient temperature during offloading, Unit Load Device (ULD) buildup, storage, and loading on the tarmacs. When ambient temperatures are too high for these critical points, the drug is easily spoiled. This prompts the need to use advanced analytics and deep learning techniques to more accurately predict temperatures around these pain points, in order to facilitate decision-making on packaging throughout the years.

DATASET DESCRIPTION

The data source used in this project was provided by the instructing team beforehand. Data contains detailed historical weather data from two weather forecasting platforms, MeteoBlue (https://www.meteoblue.com/) and OpenWeatherMap (https://openweathermap.org). Lane information was also provided as .json files.

MeteoBlue data contains both versions of high resolution and low resolution. An 'extract_city' function was used to extract historical data for each airport separately, and replaced with any information from the low resolution with high resolution. The extracted historical data for each airport was then stored in the 'combined' folder. There are 11 airports in the MeteoBlue dataset, with 41 weather variables (see Appendix 1.1.) Details of variable description can be found at MeteoBlue Weather Variables (Weather Variables · Technical Documentation). Air temperature variable used for prediction was defined as either "temperature sfc" or temperature 2 meters above the ground, depending on data availability.

The OpenWeatherMap dataset contains 5 airports information and 15 weather variables (see Appendix 1.2.) The temperature historical data of Amsterdam airport was recorded in Kelvin (K). We subtracted 272.15 from the original number to get the temperature in Celsius degree.

METHODOLOGY

Data cleaning and EDA

Observations in our datasets are organized on a daily - hourly level. Time series predictions for temperature require a fine balance between the granularity of the data and overall accuracy. In other words, the larger the prediction time frame (monthly versus weekly) the more accurate the results will be. However, given the temperature sensitivity and importance of pharmaceutical products, a level of granularity is required for practicality.

The client informed our team that they generally make packaging decisions on a monthly basis. Thus, our team decided that making hourly predictions for each month was 1) the most practical balance between accuracy and granularity 2) fit the business needs of our client, and 3) computationally possible given the scope of our computing capabilities. We averaged our predictions on an hourly monthly basis as well for consistency. Any predictor with NA values was taken out of our analysis.

After identifying our time metrics for the project, our next step was to aggregate our data and identify visually apparent patterns. Specifically, we were looking to assess yearly patterns and long term trends as our data goes back until the 1980s. After graphing temperature for each of the target cities, we found two clear trends: 1) patterns within individual years are stable and 2) temperatures in Europe have been steadily increasing overall. The latter would be a significant factor in both model selection and results.

Data Preprocessing

For each city, we extract the hourly information from 2009 to 2020. We throw away the data in 2021 because we only have the temperatures for the first half of that year which is incomplete. We consider the data between 2009 and 2016 to be our training set, the data between 2017 and 2019 to be our validation set and the data in 2020 as our test set. Also for the OWM dataset, since we have missing values in select columns, we simply drop those columns to remain consistent.

For each city, we incorporate the cyclic information of 24 hours and 12 months using trigonometric functions such as sin and cos for indicating seasonality. Since we want to predict the temperatures for a whole year ahead on a monthly basis, we aggregate our data by computing the average value of every variable for every hour in every month.

Variable Selection

After processing our data, our next step was to complete feature selections to identify which variables are important to us for predicting temperatures. Since we have two different datasets and each dataset contains different exogenous variables, we select features separately for each dataset. We created lag variables for all the variables except for the outcome (temperature) and we predict the temperature on a specific date in the current year based on other exogenous lag variables on the same date in the previous year. We made our predictions based on random forest and we ranked the feature importance utilizing the gini index.

For the Meteoblue dataset, we find some important features such as solar radiation and vapor pressure. And for the OWM dataset, we have humidity, pressure and wind speed as the significant factors.

Time Series Modeling

Autoregressive Model

For the modeling part, we first use a naïve autoregressive model. We predict the temperatures of a whole year based on the moving average of the temperatures and trigonometric values of the previous year. For simplicity, we did add any exogenous variables to the models. We import the linear regression function from sklearn to fit our model.

LSTM

As previously noted, European cities in our dataset have appeared to have consistent year over year temperature increases. When selecting an appropriate method for this project, our team was primarily concerned with this aspect of the data. We needed a method that could handle the "local" patterns within the year and the long term temperature increases. As such, we decided to use Long Short Term Memory (LSTM) neural networks. As their name suggests, LSTM models have the ability to assess both local and global patterns in the data. (Hochreiter and Schmidhuber, 1997) Thus, out of various neural network methods, we hypothesized that LSTMs would generate the most efficacious results.

For model structure we employed various strategies to make bespoke models for each of the 15 cities.

We first normalize our variables using a robust scaler to reduce the influence of the outliers. Then we tune our hyperparameters to find the optimal numbers of the LSTM layers and the optimal numbers of neurons in each layer using random search. We also add a dropout layer between the last LSTM layer and the final dense forward layer with the dropout rate ranging from 0 to 0.5 to avoid the problem of overfitting. Then we train our tuned models with the train and validation set from 10 to 20 epochs.

The model structures for the cities are as follows:

Dubai, Frankfurt, Zurich:
5 LSTM layers (1000, 800, 600, 400, 300 units)
1 dense layer with 288 units
optimizer: adam
loss function: MSE
20 epochs

Amsterdam, Basel, Brussels, Chicago, Johannesburg, Hyderabad, Lima, Narita, San Francisco, Shanghai, Santiago, Sydney:
3 LSTM layers (1000, 400, 300 units)
1 dense layer with 288 units
optimizer: adam
loss function: MSE
10 epochs

RESULTS

Metrics

After compiling our models, we then moved to assess whether the autoregressive or LSTM models performed better. At the start of the project, we utilized Mean Average Percent Error (MAPE) as our guiding metric. As the metric is an average over many percentages, it is unitless. MAPE is the standard metric in Supply Chain analyses, thus its use has a high precedent in this space. However, while MAPE is the standard in this field, our team encountered obstacles when utilizing this loss formula.

Our first obstacle was the scale of units. Celsius is not as fine of a temperature metric as Farenhight or Kelvin. For example, a five degree difference in Celcius is larger than a five degree difference in Fahrenheit. Further, many cities experience relatively low Celcius temperatures for a large part of the year. It is common for parts of Europe and the Midwest United States to see temperatures below 10 degrees Celsius during the winter. Having such a low baseline ground truth inflated the MAPE during these months. A predicted temperature of seven degrees when the ground truth is five would produce a MAPE of 40 percent. However, the same two degree difference with a ground truth of 10 degrees would only produce a MAPE of 20 percent. Thus, MAPE is not necessarily the best guiding metric when working in Celcius and in areas with colder weather.

Second, MAPE is unitless. While a percentage error may be useful when analyzing product inventory, its utility decreases when predicting temperatures with the purpose of keeping drugs safe during transportation. As previously noted, Celsius is not a fine-grain metric. A difference of just five degrees Celsius is a large temperature increase or decrease. Handling drugs, specifically vaccines, requires the most accurate temperature predictions and necessitates more specific loss measurements. To bring in units, our team decided to assess Mean Squared Error (MSE) in our analysis. MSE provides us a unit based metric and is a common error metric used across all industries. When comparing our LSTM and autoregressive models, we still took into account MAPE but weighed our decisions on which to use more on MSE. Again, this decision was made to be on the more connservative side given the drug products our client handles.

MeteoBlue

For the cities in the Meteoblue dataset, JNB, HYD, FRA, and DXB performed better with an LSTM model. This was not surprising as these cities, with the exception of FRA, had relatively stable year over year temperature trends. The temperature in FRA has been steadily increasing since at least the 1980s. The MSEs for these cities were primarily below five degrees celsius. Again, given the nature of Celcius we wanted as low of an MSE as possible. The LSTMs were great at predicting temperature during periods of steady increases and decreases but struggled in periods of temperature direction shifts.

OpenWeatherMap

For the cities in the OWM dataset, the autoregressive model performs better than the LSTM model in every city. When considering MSE as our metrics, we found out the MSEs in the autoregressive model are smaller than the MSEs in the LSTM models for every city. When it comes to MAPE as our measurement, the LSTM model only beats the autoregressive model in the case of Amsterdam. So overall the performance of the autoregressive model is better in this dataset.

PREDICTION VISUALIZATION

The final step for our project is to present our temperature prediction results to be visually attractive and straightforward to facilitate the packaging decision-making. A dashboard is developed to ensure that the results are easily interpretable. The design of the dashboard should consider the volume of information available on a single view, the functionality, as well as the visual outlook. Therefore, Tableau was used as the dashboarding tool due to its easy navigation and user-friendliness. It does not require the developer to select packages.

Due to the time constraint of the project, we developed a prototype dashboard representing the global transportation lane: Amsterdam – Frankfurt – Johannesburg (AMS-FRA-JNB).

Dashboard Construction

The csv file of predicted ambient temperatures for AMS-FRA-JNB was imported to Tableau as the data source. Then, for each of the airports on the lane, a worksheet was constructed. Temperature predictions were filtered by month and hours at the airport. The dashboard was then constructed by a grid of 3 columns. The x-axis represent the hours at the airport; specific hour ranges were hand-calculated using lane information from lane_ams-fra-jnb_lh_dtd_20210325153014 .json file. The origin was set as 12 AM representing the start of transportation. The cargo then arrives at Amsterdam Airport Schiphol at 3AM, and leaves at 9AM. For Frankfurt International Airport, it arrives at 5PM and leaves at 4:30PM the next day. The cargo then arrives at O. R. Tambo International Airport at 2:55AM and leaves at 8:55AM on day 3. The y-axis represents the average predicted ambient temperatures at each airport in the specific hour ranges throughout the month that the user specifies. On the upperight, the user can choose the month that she/he is interested in looking at. Thus, by inputting the month number, the user can get hourly ambient temperature predictions on each airport across the entire lane throughout the day.

To facilitate decision-making for different product temperatures, 3 different possible product temperature ranges were illustrated using boxes in different colors on the dashboard. The blue box represents product temperature ranges from 15 °C to 25 °C; yellow box represents product temperature ranges of 2 °C to 8 °C; green box represents product temperature ranges between -20°C and -10°C. The user can then determine whether the ambient temperature predictions are in the range of product temperature by looking at whether the dots fall out of the boxes, to make packaging decisions. For example, at Amsterdam Airport Schiphol in January 2020, for a product range of 2 °C - 8 °C, the ambient temperature would be suitable for a softbox storage scheme; however, for a product range from -20°C to -10°C, an active cooling container should be utilized. (Figure 1)

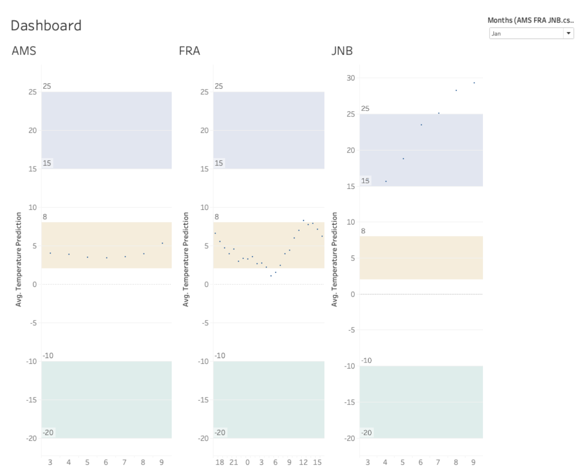

Figure 1. Ambient temperature predictions for AMS-FRA-JNB in January 2020.

Future Work/Conclusions

In terms of the machine learning life cycle, a yearly update of our models will likely be necessary. Previously mentioned uncertainty related to future temperature trends requires that the models be trained on the most recent data. In areas in which year over year temperature does not change significantly (e.g. Dubai) the company may be able to get away with lengthening the periods between model updates. However, for optimal accuracy we recommend a yearly update to ensure that packaging decisions incorporate the most recent data.

Accurate temperature predictions are a vital part of the pharmaceutical supply chain process. The sensitivity of drugs necessitates that any prediction model has a small margin of error. Through our analysis we explored the use of LSTM neural networks versus more traditional autoregressive methods in forecasting temperature across a variety of freight lanes. LSTM models generally performed better at capturing stable year over year trends. Autoregressive models performed better with cities that have experienced consistent year over year temperature increases. The ability to predict temperatures in areas most affected by climate change will likely become harder given the uncertainty of how future trends will deviate from previous data. Thus, relying on multiple model strategies is paramount for success. A combination of LSTM and autoregressive models will give Validaide a strong set of tools necessary to successfully complete their pharmaceutical missions.

Citations

1. The rules for shipping pharmaceuticals you need to know. ABCO Transportation. Nov. 12, 2018. https://www.shipabco.com/the-rules-for-shipping-pharmaceuticals-you-need-to-know/#:~:text=Transportation%3A%20The%20medicines%20or%20pharma,inventory%20at%20the%20correct%20temperature.
2. Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." Neural computation 9.8 (1997): 1735-1780.

Appendix

1.1 Weather Variables in MeteoBlue Dataset

Temperature 2 m elevation corrected
Growing Degree Days 2 m elevation corrected
Temperature 1000 mb
Temperature 850 mb
Temperature 700 mb
Precipitation Total sfc
Snowfall Amount sfc
Relative Humidity 2 m
Wind Speed 10 m
Wind Direction 10 m
Wind Speed 80 m
Wind Direction 80 m
Wind Gust sfc
Wind Speed 900 mb
Wind Direction 900 mb
Wind Speed 850 mb
Wind Direction 850 mb
Wind Speed 700 mb
Wind Direction 700 mb
Wind Speed 500 mb
Wind Direction 500 mb
Cloud Cover Total sfc
Cloud Cover High high cld lay
Cloud Cover Medium mid cld lay
Cloud Cover Low low cld lay
CAPE 180-0 mb above gnd
Sunshine Duration sfc
Shortwave Radiation sfc
Direct Shortwave Radiation sfc
Diffuse Shortwave Radiation sfc
Mean Sea Level Pressure MSL
Geopotential Height 1000 mb
Geopotential Height 850 mb
Geopotential Height 700 mb
Geopotential Height 500 mb
Evapotranspiration sfc
FAO Reference Evapotranspiration 2 m
Soil Temperature 0-10 cm down
Soil Moisture 0-10 cm down
Vapor Pressure Deficit 2 m

1.2 Weather Variables in OpenWeatherMap Dataset

Temperature
feels_like
temp_min
temp_max
pressure
sea_level
grnd_level
Humidity
wind_speed
wind_deg
rain_1h
rain_3h
snow_1h
snow_3h
clouds_all