Food Prices

Final Project

Data Analytics Boot Camp '19

Blanca Chavarría | Nancy Torres | Rubén Colmenares | Yozgart Álvarez

Overview

- How much does basic food products cost?
- Which countries produces them?
- How much do the prices change over time?

These are some questions we posed for our project and inspired us for the following analysis.

Datasets

- FAO, yearly North America prices from 2000 to 2017 – Mexico's Trade from 1986 to 2016 – Mexico's Production from 1961 to 2017
- INEGI, monthly average Mexico prices for 2011 - 2017

Approach

Technologies used for the project

- Python
- Pandas
- Matplotlib
- Scikit-Learn
- Statsmodels
- HTML / CSS / Bootstrap
- JavaScript
- Amazon AWS
- Tableau

Dashboard

Tomato price prediction with ARIMA

- **Goal**. Predict the price per kg. of tomato (saladette) in CDMX
- **Data**. Prices by INEGI from 2011 to 2017
- **Algorithm**. ARIMA as it is meant for time serie, requires more than 50 observations that is why information is per month
- **Results**. An ARIMA model that can predict prices for future months for the short term

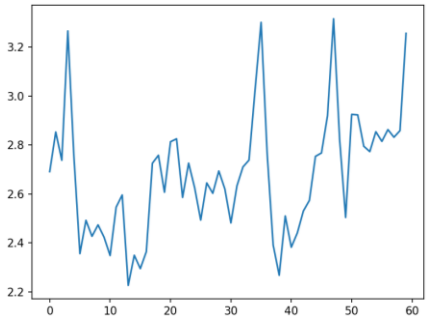
ARIMA Model Results

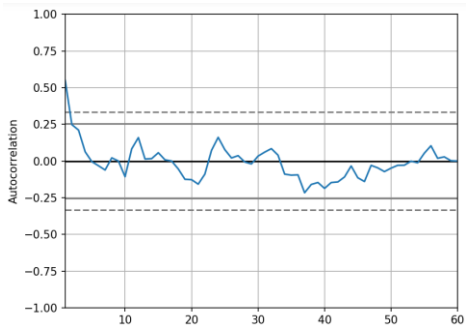
Dep. Variable:	D.Precio promedio	No. Observations:	58
Model:	ARIMA(3, 1, 1)	Log Likelihood	-161.293
Method:	css-mle	S.D. of innovations	3.730
Date:	Fri, 09 Aug 2019	AIC	334.586
Time:	14:09:08	BIC	346.949
Sample:	1	HQIC	339.402

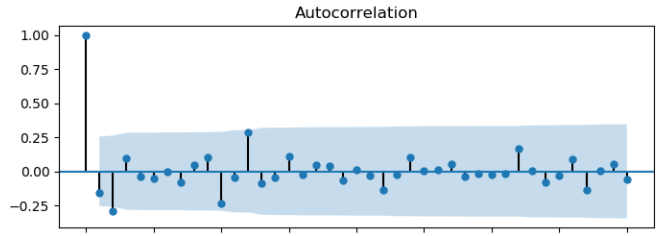
	coef	std err	z	P>|z|	[0.025	0.975]
const	0.0151	0.018	0.825	0.413	-0.021	0.051
ar.L1.D.Precio promedio	-0.2326	0.142	-1.638	0.107	-0.511	0.046
ar.L2.D.Precio promedio	-0.3430	0.131	-2.622	0.011	-0.599	-0.087
ar.L3.D.Precio promedio	-0.0320	0.150	-0.214	0.831	-0.325	0.261
ma.L1.D.Precio promedio	-0.9999	0.047	-21.377	0.000	-1.092	-0.908

Roots

	Real	Imaginary	Modulus	Frequency
AR.1	-0.2054	-1.7289j	1.7411	-0.2688
AR.2	-0.2054	+1.7289j	1.7411	0.2688
AR.3	-10.2977	-0.0000j	10.2977	-0.5000
MA.1	1.0001	+0.0000j	1.0001	0.0000

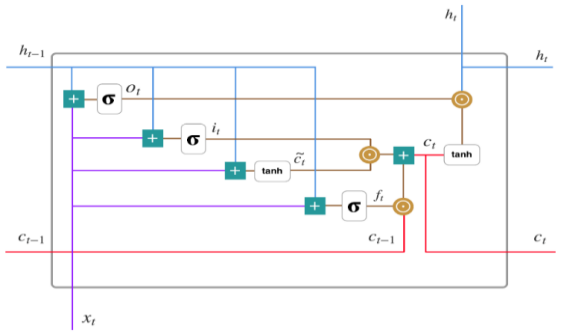






Tomato price prediction with LSTM

- **Goal**. Predict whether prices will be going up or going down using Machine Learning
- **Data**. Prices reported by Mexico to the FAO from 1991 to 2017
- **Algorithm**. LSTM are well suited to learn from important experiences that have very long-time lags in between; includes memory states that decide what information should be propagated further at each timestep
- **Results**. On average the **Hyper tuned Bidirectional model using GridSearchCV** was better (the difference between the actual and the predicted values was smaller)

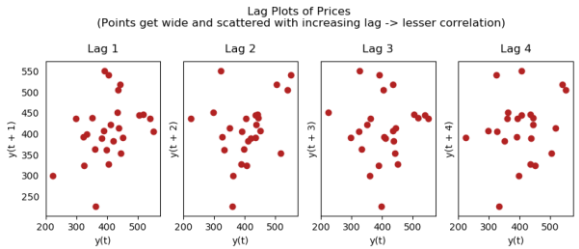


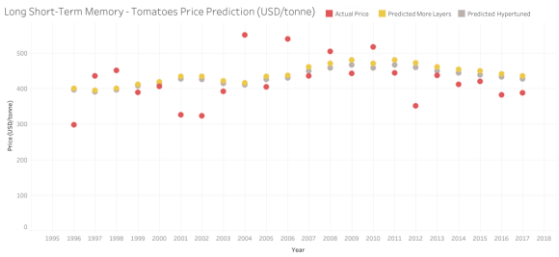
o_t is the output gate

i_t is called the input gate

$\tilde{c}_t$ is the value for the new memory cell

f_t is the forget gate





+1 23 987 6554