



Sistema de Recomendación de Inversión Bursátil en base a Predicciones de Mercado y Análisis de Opinión de Twitter

Álvaro Bartolomé del Canto



VNIVERSIDAD
D SALAMANCA

CAMPUS DE EXCELENCIA INTERNACIONAL

Introducción

1. Introducción
2. Características Principales
3. Resultados
4. Trabajo Futuro

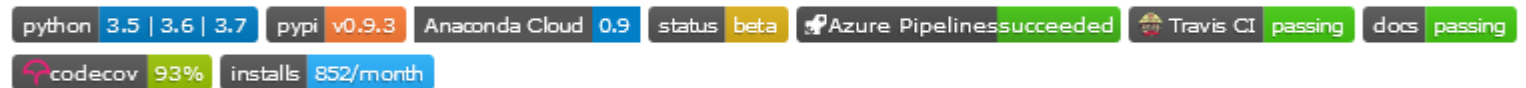
- El proyecto consiste en **predecir el comportamiento futuro de las acciones del mercado continuo español** a partir de los datos recuperados de Investing.com.
- Para la predicción se propone la combinación de técnicas de **análisis técnico financiero** como técnica de predicción tradicional, combinada junto con **algoritmos de Machine Learning** de regresión.
- El sistema, por tanto, funciona como un **recomendador** que, a partir de las predicciones, busca la **identificación de señales de compra o venta** de acciones en el mercado.
- Adicionalmente, con el fin de apoyar los datos históricos del valor diario de las acciones, se propone la adición de la **opinión de Twitter** sobre una acción con el fin de determinar el impacto positivo o negativo sobre el comportamiento de la misma.

1. Introducción
2. Características Principales
3. Resultados
4. Trabajo Futuro

Características

- La plataforma creada para visualizar tanto los datos como la predicción y recomendación de la acción requiere la ingesta de datos de Investing.com, la fuente fiable elegida para dicho fin.
- De este modo, se ha creado un paquete de Python llamado **investpy** para encapsular la recuperación de los datos históricos de Investing.com para las acciones del mercado continuo español.
- Adicionalmente, **investpy** es escalable y cubre la recuperación de información tanto de acciones, fondos y ETFs de todo el mundo.

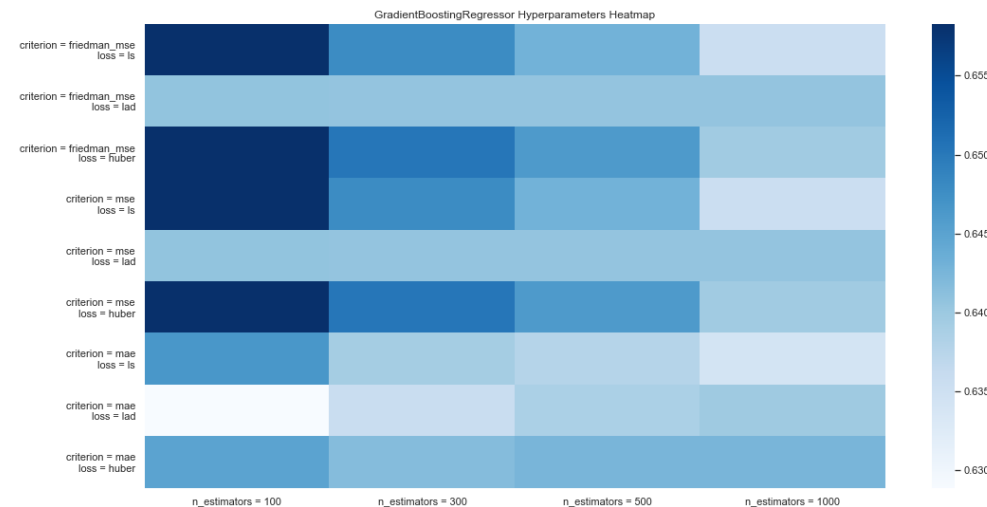
investpy — a Python package for financial historical data extraction from Investing



Características

1. Introducción
2. Características Principales
3. Resultados
4. Trabajo Futuro

- Por tanto, una vez desarrollado el sistema de extracción de datos históricos de las acciones del mercado continuo español, se procederá al análisis de los mismos.
- Junto con el **análisis de datos** se incluyen el estudio y las pruebas de los distintos **algoritmos de Machine Learning** a utilizar, en este caso, los algoritmos de regresión en los que se predice el valor de una variable objetivo en base a otro valor (se predice el valor de cierre en función del valor de apertura).
- El estudio de los algoritmos incluye la validación cruzada de los hiper-parámetros de estos, con el fin de determinar la mejor combinación de hiper-parámetros para cada algoritmo y, a su vez, para cada acción.



Características

1. Introducción
2. Características Principales
3. Resultados
4. Trabajo Futuro

- Para comenzar el estudio planteado en el proyecto con objetivo de determinar el impacto del sentimiento analizado en base a la opinión de Twitter para con el comportamiento del mercado (tendencias), se propone la creación de un paquete de Python llamado **trendet** que dada una serie temporal, detecta todas las tendencias encontradas, tanto alcistas como bajistas.
- **trendet** es un paquete orientado a la detección de tendencias en series temporales de datos históricos financieros, de modo que está integrado y pensado para su uso junto con **investpy**.

trendet - is a Python package for trend detection on stock time series data

python 3.5 | 3.6 | 3.7 | pypi v0.5 | status alpha | Azure Pipelines succeeded | Travis CI passing | docs passing | codecov 97%
downloads 336/month



trendet is a Python package for trend detection on stock time series data

Características

1. Introducción
2. Características Principales
3. Resultados
4. Trabajo Futuro

- De este modo, con el fin de recuperar tweets dada una tendencia de una serie temporal entre dos fechas para, más adelante, analizar el sentimiento de dichos tweets, se propone la creación de un paquete de Python para la ingesta de tweets con el API de Twitter tanto en batch como en streaming, para las versiones Free y Premium de la API.
- La creación de **twipper** surge a modo de solventar las carencias del resto de paquetes/bibliotecas de Python para usar la API de Twitter, combinando ambos planes ofertados por Twitter.

twipper - Twitter Wrapper written in Python

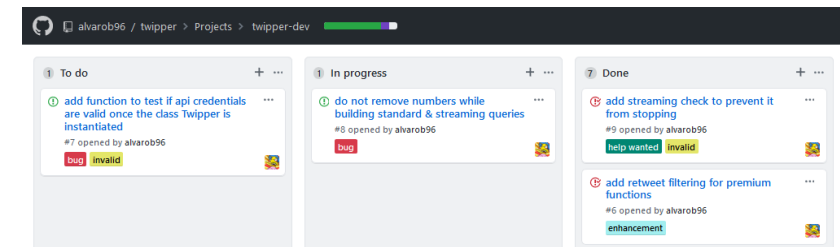
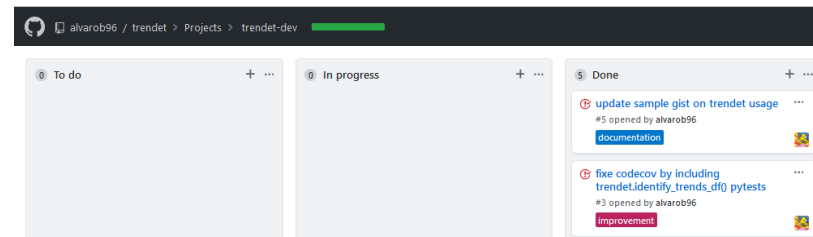
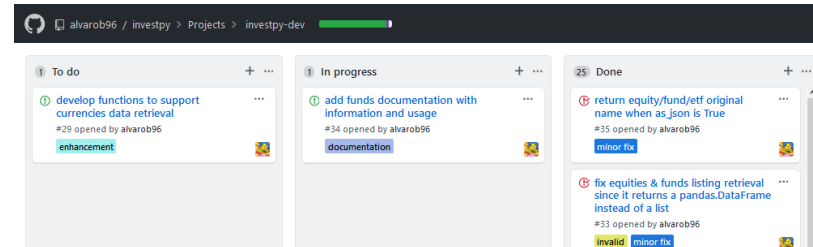
python 3.5 | 3.6 | 3.7 | pypi v0.1.3 | status alpha | Azure Pipelines succeeded | Travis CI passing | docs passing | downloads 265/month



Características

- Adicionalmente, como resultado de los paquetes de Python creados, concluir que se ha seguido un desarrollo basado en **Kanban Automático**, auto-gestionado a través de **GitHub Projects**. Esta utilidad integrada con GitHub permite gestionar las tareas o issues del proyecto en una tabla de Kanban con desplazamiento automático de las tareas, lo cual facilita enormemente las labores de gestión.

1. Introducción
2. Características Principales
3. Resultados
4. Trabajo Futuro



Características

1. Introducción
2. Características Principales
3. Resultados
4. Trabajo Futuro

- Como resultado de dicho análisis, se crea una plataforma para la visualización y recomendación de inversión bursátil en base a la combinación del análisis técnico con los algoritmos de Machine Learning.
- Adicionalmente, para la clasificación de tweets ingeridos en tiempo real se crea una plataforma para la clasificación de los mismos de forma supervisada y mecánica en sentimientos, siendo estos: **positivo, negativo y neutral**.



Plataforma de Visualización y Recomendación de Inversión Bursátil en Acciones del Mercado Continuo Español

Visualización

Visualización y Recomendación

Álvaro Bartolomé del Canto, 2018-2019 @ alvarob96 en GitHub



Plataforma para la Clasificación Supervisada de Tweets en Sentimientos
Recuperados con twipper

Clasificador de Sentimientos

Álvaro Bartolomé del Canto, 2018-2019 @ alvarob96 en GitHub

- 1. Introducción
- 2. Características Principales
- 3. Resultados
- 4. Trabajo Futuro

Resultados

- La plataforma resultante para la predicción del comportamiento futuro de las acciones a través de los algoritmos de Machine Learning resulta en un gráfico de barras con la precisión de los algoritmos aplicados:



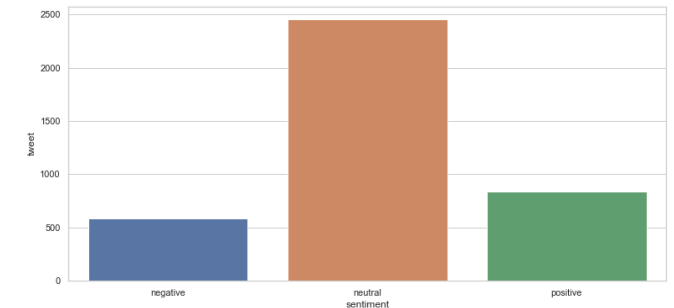
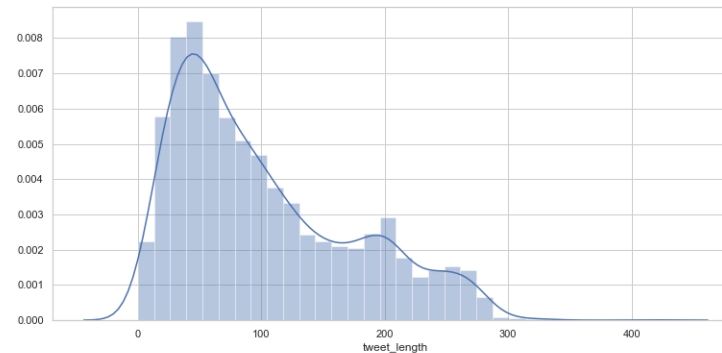
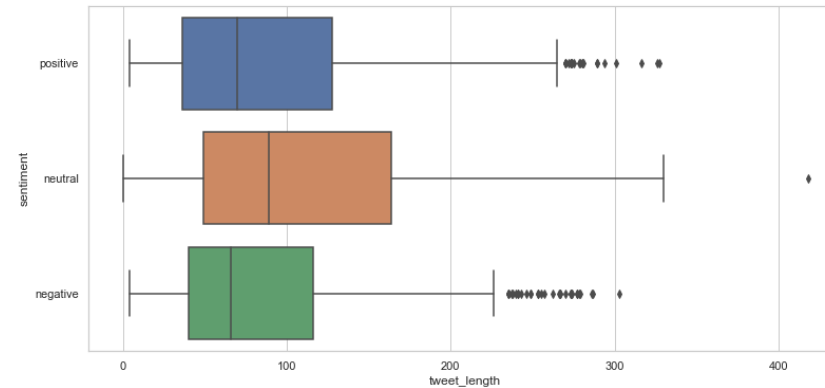
- Además, la recomendación resultante de los distintos algoritmos y técnicas aplicadas se visualiza en una tabla junto con su justificación en forma de señales de compra/venta. Por ejemplo:

Factor Técnico	Resultado	Señal
RSI - 14 días	63.19919	COMPRAR
STOCH - 9,6 días	76.69109	COMPRA FUERTE
ULTOSC - 7,14,28 días	64.36916	COMPRAR
WILLR - 14 días	-12.85714	SOBRECOMPRA

Resultados

1. Introducción
2. Características Principales
3. Resultados
4. Trabajo Futuro

- Los resultados obtenidos de la clasificación supervisada de tweets en sentimientos son los mostrados a continuación, que reflejan el groso de datos recuperados y etiquetados en sentimientos **positivo, negativo y neutral**.

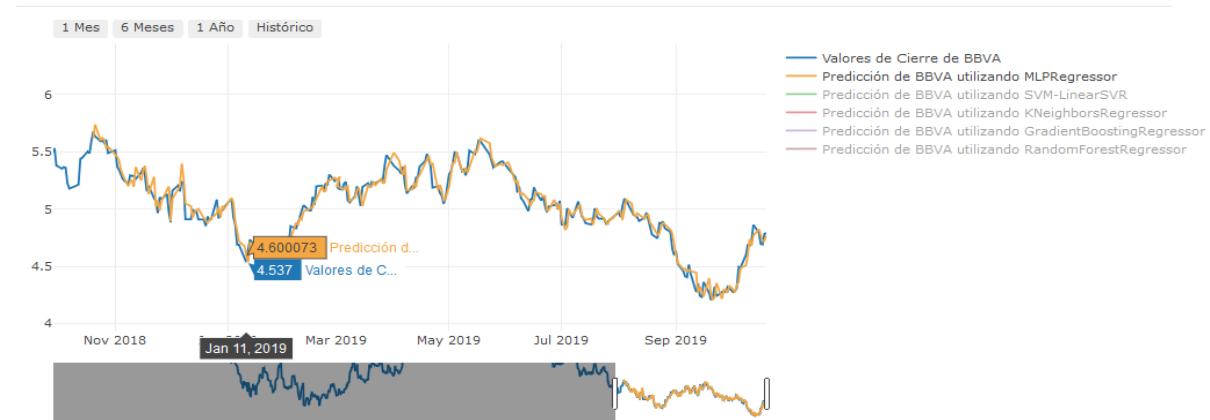


Resultados

1. Introducción
2. Características Principales
3. Resultados
4. Trabajo Futuro

- Finalmente, concluir que la combinación entre ambas técnicas de análisis técnico e Inteligencia Artificial, aunque no siempre similar, las predicciones eran bastante parecidas a la hora de identificar señales de compra/venta, con la salvedad de que la predicción a través de los algoritmos de Machine Learning era tan solo aplicable al siguiente día hábil.

Predicción de BBVA utilizando Algoritmos de Regresión



- El impacto de la opinión por medio del análisis de sentimiento de Twitter junto con la identificación de las tendencias pasadas a modo de caso de estudio no es concluyente. Esto se debe a las grandes carencias que hay de soluciones para el NLP en español, y más en concreto, en lo que a análisis de sentimientos se refiere por la ausencia de datasets clasificados en español. Esto implica que aún se encuentra bajo proceso de investigación y desarrollo.

Trabajo Futuro

1. Introducción
2. Características Principales
3. Resultados
4. Trabajo Futuro

- Mejorar el modelo de clasificación no supervisada de tweets en sentimientos ya que su actual precisión es tan solo del 70% para un dataset de 4000 tweets.
- Ampliar las funcionalidades tanto de **investpy**, como de **trendet** y de **twipper**; con el fin de cubrir todas las necesidades existentes o futuras.
- Ampliar la predicción y recomendación a acciones de todo el mundo (o por lo menos de más países).
- Investigar sobre el posible encadenamiento de predicciones sucesivas de cara a predecir el comportamiento futuro con Inteligencia Artificial más allá de un día.
- Investigar y analizar el impacto de Twitter en el comportamiento del mercado una vez que el modelo tenga más precisión, o crear un modelo para predecir sentimiento con vocabulario específico orientado al mercado de acciones.



¡Gracias!

Contacto

✉ alvarob96@usal.es

🐙 @alvarob96



VNiVERSIDAD
D SALAMANCA

CAMPUS DE EXCELENCIA INTERNACIONAL