Universidad Central del Ecuador

FAC U LTAD D E I N G E N I E R ÍA Y

C I E N C IAS APLI CAD AS

**Sistemas De Información**

Documento de Selección del Modelo



**Estudiante:**

* Luis Angel Gaona Cumbicus

[lagaona@uce.edu.ec](mailto:lagaona@uce.edu.ec)

* Raul Alexander Pazos Erraez

[rapazos@uce.edu.ec](mailto:rapazos@uce.edu.ec)

* Cristian Daniel Toca Rocha

[cdtoca@uce.edu.ec](mailto:cdtoca@uce.edu.ec)

* Marlon Josue Espinosa Mancero

mjespinosam@uce.edu.ec

**Docente:**

* PhD, Jefferson Tarcisio Beltrán Morales

[jtbeltran@uce.edu.ec](mailto:jtbeltran@uce.edu.ec)

**Asignatura:** Minería de datos

**Paralelo:** S8-P2

**Fecha:** sábado 13 de julio de 2024

Logotipo

Descripción generada automáticamente

**Generative AI-Powered Economic Impact Analysis System (GEIA)**

***Fecha:****13/07/2024*

Contenido

[HOJA DE CONTROL 4](#_Toc174713229)

[Historial de Cambios 4](#_Toc174713230)

[Introducción 5](#_Toc174713231)

[Modelos Considerados 5](#_Toc174713232)

[1. Prophet 5](#_Toc174713233)

[2. CausalImpact 7](#_Toc174713234)

[TIPO DE MODELOS 7](#_Toc174713235)

[Resultados y Justificación 14](#_Toc174713236)

HOJA DE CONTROL

|  |  |  |  |
| --- | --- | --- | --- |
| Organismo | Universidad Central Del Ecuador | | |
| Proyecto | Generative AI-Powered Economic Impact Analysis System (**GEIA**) | | |
| Entregable | Documento de selección del modelo | | |
| Autor | Luis Angel Gaona Cumbicus | | |
| Versión/Edición | V1.0 | Fecha Versión | 13/07/2024 |
| Aprobado por |  | Fecha Aprobación | .…./…../….. |
|  |  | N.º Total de Páginas | 15 |

Historial de Cambios

|  |  |  |  |
| --- | --- | --- | --- |
| Fecha | Autor | Organización | Descripción |
|  |  |  |  |

Introducción

En el proyecto de implementación de un modelo predictivo para estimar los daños económicos en diferentes sectores de la economía ecuatoriana causados por eventos adversos, se requiere seleccionar un modelo adecuado para el análisis de series de tiempo interrumpidas. Para este propósito, se comparan dos modelos prominentes: **Prophet y CausalImpact.**

Modelos Considerados

1. Prophet

Prophet es un modelo desarrollado por Facebook para la predicción de series de tiempo que manejan estacionalidad diaria, semanal y anual. Es especialmente útil para datos con tendencias estacionales y cambios abruptos. Prophet modela las tendencias y estacionalidades con componentes aditivos y puede manejar cambios en la estacionalidad a lo largo del tiempo.

Texto

Descripción generada automáticamente

Texto

Descripción generada automáticamente

Gráfico, Gráfico de líneas, Histograma

Descripción generada automáticamente

2. CausalImpact

CausalImpact es una herramienta de Google para la evaluación del impacto de intervenciones en series de tiempo. Utiliza un modelo bayesiano para estimar el impacto de un evento en una serie temporal comparando el comportamiento real con un contrafactual, es decir, cómo se habría comportado la serie temporal en ausencia del evento.

TIPO DE MODELOS

1. *Texto

   Descripción generada automáticamenteTexto

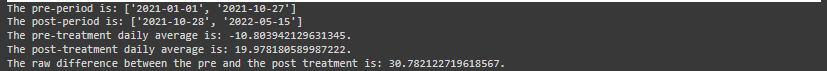
   Descripción generada automáticamente****Modelo univariable***

Interfaz de usuario gráfica, Gráfico, Histograma

Descripción generada automáticamente Forma

Descripción generada automáticamente con confianza media

Gráfico

Descripción generada automáticamente

Texto

Descripción generada automáticamente *Durante el periodo post-intervención, la variable respuesta tuvo  
un valor medio de aprox. 19,98. Por el contrario, en ausencia de una  
intervención, hubiéramos esperado una respuesta promedio de -6,54.  
El intervalo del 95% de esta predicción contrafactual es [-8,83, -4,11].  
Restando esta predicción de la respuesta observada se obtiene  
una estimación del efecto causal que tuvo la intervención sobre la  
variable de respuesta. Este efecto es 26,52 con un intervalo del 95% de  
[24.09, 28.81]. Para una discusión sobre la importancia de este efecto,  
vea abajo.  
  
Resumir los puntos de datos individuales durante la posintervención.  
período (que sólo a veces puede ser interpretado significativamente), el  
La variable respuesta tuvo un valor global de 3995,64.  
Por el contrario, si la intervención no hubiera tenido lugar, habríamos esperado  
una suma de -1308,33. El intervalo del 95% de esta predicción es [-1765,94, -822,13].  
  
Los resultados anteriores se dan en términos de números absolutos. En relativo  
En términos generales, la variable respuesta presentó una disminución del -405,4%. el 95%  
El intervalo de este porcentaje es [-440,38%, -368,24%].  
  
Esto significa que el efecto negativo observado durante la intervención  
período es estadísticamente significativo.  
Si el experimentador esperaba un efecto positivo, se recomienda  
para verificar si las anomalías en las variables de control pueden tener  
provocó una expectativa demasiado optimista de lo que debería haber sucedido  
en la variable respuesta en ausencia de la intervención.  
  
La probabilidad de obtener este efecto por casualidad es muy pequeña.  
(Probabilidad bayesiana del área de la cola unilateral p = 0,0).  
Esto significa que el efecto causal puede considerarse estadísticamente.  
significativo.*

1. Texto

   Descripción generada automáticamenteTexto

   Descripción generada automáticamente***Modelo multivariable***

Texto

Descripción generada automáticamente

Texto

Descripción generada automáticamenteGráfico

Descripción generada automáticamente con confianza media

*Durante el periodo post-intervención, la variable respuesta tuvo  
un valor medio de aprox. 30.08. Por el contrario, en ausencia de una  
intervención, hubiéramos esperado una respuesta promedio de 40,03.  
El intervalo del 95% de esta predicción contrafáctica es [37,81, 42,41].  
Restando esta predicción de la respuesta observada se obtiene  
una estimación del efecto causal que tuvo la intervención sobre la  
variable de respuesta. Este efecto es -9,94 con un intervalo del 95% de  
[-12,33, -7,72]. Para una discusión sobre la importancia de este efecto,  
vea abajo.  
  
Resumir los puntos de datos individuales durante la posintervención.  
período (que sólo a veces puede ser interpretado significativamente), el  
La variable respuesta tuvo un valor global de 6016,92.  
Por el contrario, si la intervención no hubiera tenido lugar, habríamos esperado  
una suma de 8005,59. El intervalo del 95% de esta predicción es [7561,42, 8482,55].  
  
Los resultados anteriores se dan en términos de números absolutos. En relativo  
En términos generales, la variable respuesta presentó una disminución del -24,84%. el 95%  
El intervalo de este porcentaje es [-30,8%, -19,29%].  
  
Esto significa que el efecto negativo observado durante la intervención  
período es estadísticamente significativo.  
Si el experimentador esperaba un efecto positivo, se recomienda  
para verificar si las anomalías en las variables de control pueden tener  
provocó una expectativa demasiado optimista de lo que debería haber sucedido  
en la variable respuesta en ausencia de la intervención.  
  
La probabilidad de obtener este efecto por casualidad es muy pequeña.  
(Probabilidad bayesiana del área de la cola unilateral p = 0,0).  
Esto significa que el efecto causal puede considerarse estadísticamente.  
significativo.*

Resultados y Justificación

En este proyecto, hemos evaluado dos modelos prominentes para el análisis de series de tiempo interrumpidas: Prophet y CausalImpact. Ambos modelos tienen sus fortalezas y aplicaciones específicas.

**Prophet** es una herramienta poderosa para realizar predicciones en series temporales, especialmente en datos con tendencias estacionales y cambios abruptos. Sin embargo, Prophet está diseñado principalmente para la predicción de tendencias futuras y no se centra en la evaluación del impacto de eventos específicos en la serie temporal. Si bien es útil para modelar el comportamiento general de la serie, no ofrece una metodología directa para cuantificar cómo un evento adverso afecta la serie temporal en comparación con un escenario contrafactual (sin el evento).

Por otro lado, **CausalImpact** es un modelo diseñado específicamente para medir el impacto de un evento en una serie temporal, lo que lo hace ideal para el problema que queremos resolver en este proyecto. CausalImpact no solo permite realizar un análisis contrafactual, sino que también ofrece un reporte detallado y resumido del impacto del evento, proporcionando claridad sobre la diferencia entre el escenario con y sin el evento. Este nivel de detalle es crucial para entender cómo un evento adverso afecta los diferentes sectores de la economía ecuatoriana.

Además, CausalImpact ofrece varias ventajas adicionales, como la posibilidad de integrar información de múltiples variables explicativas y la capacidad de generar intervalos de confianza para las estimaciones del impacto, lo que mejora la robustez y la interpretabilidad de los resultados. Estas características hacen que CausalImpact sea la herramienta más adecuada para nuestro objetivo de estimar los daños económicos causados por eventos adversos.

Por estas razones, se ha decidido adoptar CausalImpact como el modelo principal para el análisis de series de tiempo interrumpidas en este proyecto. Este enfoque nos permitirá obtener una visión clara y detallada del impacto económico de eventos adversos, contribuyendo a la toma de decisiones informadas para mitigar dichos impactos en el futuro.