

Módulo IV: SISTEMA SOPORTE DE PREDICCIONES

Ma. Laura CALIUSCO¹, Pablo VILLARREAL¹, Alejandro TOFFOLO¹, Ma Laura TAVERNA¹, Omar CHIOTTI²

(1) **GIDSATD - UTN** - Facultad Regional Santa Fe - Lavaisse 610 - 3000 SANTA FE - ARGENTINA

(2) **INGAR - CONICET** - Avellaneda 3657 - 3000 Santa Fe - e-mail: chiotti@arcrde.edu.ar

RESUMEN

En este trabajo presentamos una descripción de un subsistema que brinda soporte a la actividad de predicciones. Este subsistema forma parte de un Sistema Soporte de Decisiones globales prototipo en desarrollo en el GIDSATD (Grupo de Investigación y desarrollo de Sistemas de Ayuda a la Toma de Decisiones). En particular, en primer lugar describimos brevemente la base de modelos del sistema, en segundo lugar presentamos un análisis y diseño orientado a objetos del sistema soporte de predicciones (SSP).

INTRODUCCIÓN

En un proceso de toma de decisiones es necesario considerar las variaciones de la información que pueden producirse a lo largo del horizonte de tiempo que involucra la decisión. Lorino (1989) establece... "tomar la decisión en base a los datos actuales es una simplificación frecuentemente inadmisibles del problema"... Esto plantea la necesidad de disponer de herramientas de predicción como soporte para la toma de decisión.

Por otra parte, ... "intentar racionalizar y tecnificar al extremo el proceso de toma de decisión es tanto como ignorar sus condicionantes de oportunidad, imaginación y eficacia relativa"... de este modo se plantea aquí que si bien se requieren herramientas, estas no debieran ser pensadas como reemplazantes del experto en predicciones. Como ejemplo de desarrollos que apuntan en este sentido podemos citar el trabajo de Lo (1994), el autor describe un sistema experto, el cual contiene el conocimiento necesario para seleccionar el método de predicción más apropiado para predecir una demanda. A diferencia de los sistemas expertos, un sistema soporte de decisiones es pensado como una herramienta que brinda soporte al experto y no en reemplazo del mismo. Es más, es una herramienta poco útil para alguien que no es experto.

Otro aspecto a considerar es que una predicción depende del entorno en el cual se enmarca, por lo tanto no puede ser permanente sino que debe estar sometida a un proceso continuo de revisión y perfeccionamiento. Esto plantea la necesidad de integración del SSP con otros sistemas soporte de decisiones de la organización ya que la información generada por un proceso de predicción ha de ser utilizada en otras actividades de decisión, tales como planificación estratégica, planificación a mediano y corto plazo, planificación de compras de materiales, etc. Si bien se han desarrollado softwares que pueden ser utilizados para ayudar en la tarea de predicciones, la integración de estos en un sistema es generalmente difícil y poco eficaz.

Con este rápido análisis intentamos reflejar la necesidad de disponer de un soporte global de información. Esto es, los expertos de cada área de una organización requieren de herramientas e información para tomar decisiones. Los modelos y herramientas generalmente son particulares de un área mientras que la información es global. Por lo tanto, un sistema soporte de decisiones competitivo requiere de una adecuada arquitectura y no puede surgir de la integración de softwares diseñados en forma independiente. Nuestro grupo está trabajando en el diseño de la arquitectura de un SSD global (Rico y otros, 1997). Con el término global nos referimos a un sistema que involucra todos los puntos de decisión de una organización y además que no es diseñado para una organización en particular. En el desarrollo de este sistema, una de las actividades a brindar soporte es la de predicciones. En este trabajo describimos la arquitectura del

subsistema soporte de predicciones. En primer lugar describimos brevemente la conformación de la base de modelos del SSP, luego presentamos un análisis y diseño orientado a objeto del SSP.

BASE DE MODELOS DEL SSP

La base de modelos del SSP está integrada por modelos de predicción y herramientas estadísticas necesarias para dar soporte a los procesos de predicción.

Técnicas de predicción: se pueden clasificar en función del tipo de información que utilizan en:

Técnicas de información subjetiva: utilizan la opinión que tienen ciertas personas sobre el futuro de la situación en estudio. Estas técnicas tratan de obtener el máximo aprovechamiento de las experiencias, opiniones y expectativas de personas o instituciones seleccionadas según diversos criterios. Puede ser el caso de encuestas de opinión de empresarios, expectativas de compra de ciertos consumidores, etc. En esta clasificación se incluyen los modelos de captura del entorno.

Técnicas de información histórica: se basan en la evolución en períodos anteriores del fenómeno objeto de estudio. Utilizan la serie temporal de la variable económica en estudio. Como ejemplo podemos citar: predecir la inflación del próximo trimestre en base a los datos mensuales de los últimos años, predecir las ventas de una empresa para el próximo año en base a la evolución de las mismas en años anteriores.

Técnicas de información relacional o causal: se basan en las relaciones internas de funcionamiento del tema cuyo comportamiento se trata de predecir. Utilizan la conexión entre las variables condicionantes del tema. Las leyes de comportamiento que se supone van a regir en el futuro, es frecuente que se deduzcan de la experiencia sobre su funcionamiento en el pasado, con lo cual la información histórica resulta también relevante. Como ejemplo podemos considerar la situación de predecir la evolución del tipo de cambio en base a la posible marcha de otros aspectos que la condicionan, tales como el saldo de la balanza comercial, las diferencias de las tasas de interés o de inflación entre países.

Las técnicas incorporadas a la base de modelos del SSP correspondientes a cada grupo, se representan en la Figura 1:

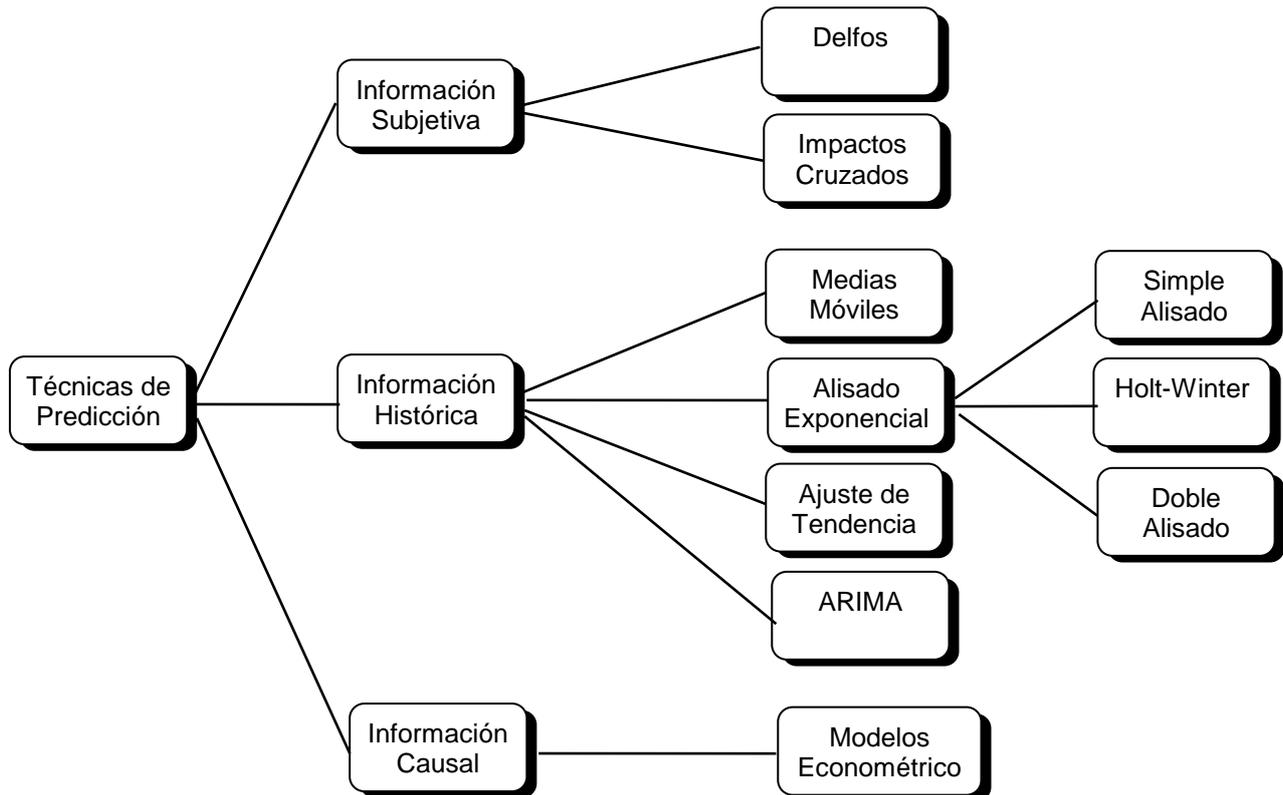


Figura 1

Herramientas Estadísticas: Los métodos estadísticos que incorporamos al módulo son: cálculo de la media, la varianza y el desvío standard. Además, el sistema permite realizar el cálculo de las frecuencias relativas y absolutas de una serie, así como también las gráficas de histograma y ojiva. Para completar el análisis estadístico agregamos distintos test de hipótesis para realizar validaciones. De ellos incluimos: test de la media, test de la varianza, test de correlación, Contraste_t y Contraste_F.

ANÁLISIS Y DISEÑO ORIENTADO A OBJETOS

En el diseño de este módulo hemos utilizado la metodología de análisis y diseño orientado a objetos propuesta por Grady Booch (1993). Para obtener los requerimientos de Análisis utilizamos el enfoque de Análisis de Casos. El mismo consiste en definir escenarios que en su conjunto describen las funciones a las que debe responder el módulo. A través del recorrido de cada escenario se identificaron los objetos que participan, las responsabilidades de cada uno, y como cada objeto colabora con otros, en el sentido de las operaciones que cada uno invoca sobre los otros. Consideremos como ejemplo el escenario: *Predicción usando Alisado Exponencial*.

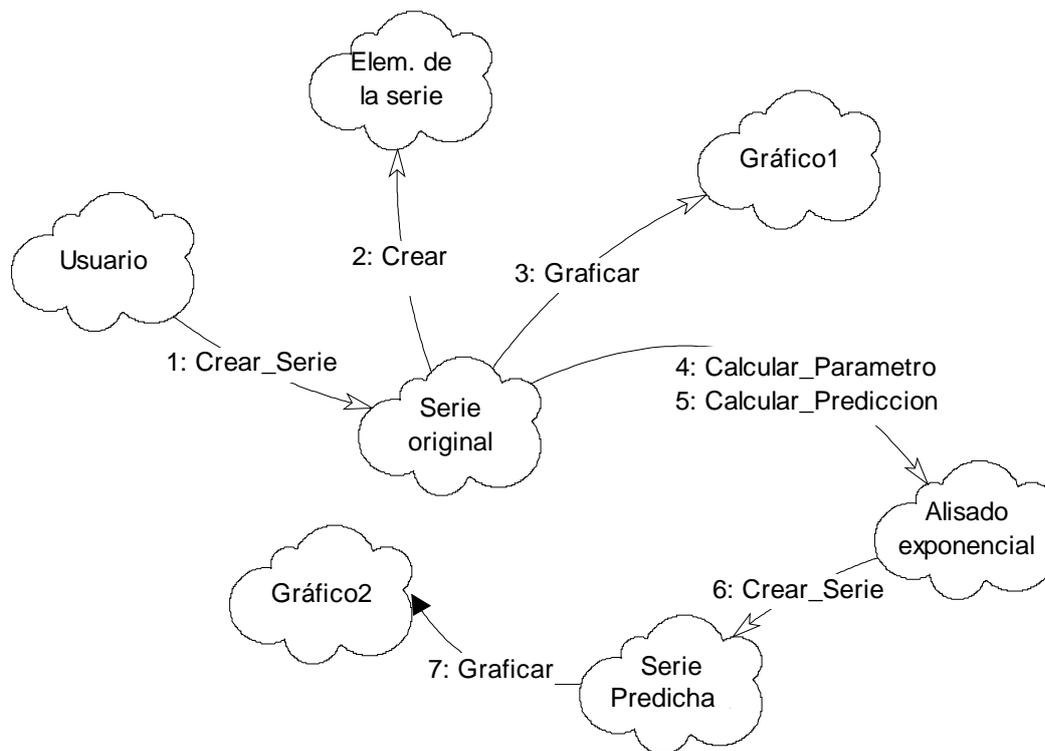
Alisado Exponencial: Se interpreta como una media ponderada de valores previos anteriores, reales y de predicción. El alisado exponencial es una técnica donde la predicción se genera mediante un mecanismo de cálculo recursivo establecido a priori.

Para el cálculo se puede utilizar la expresión $\gamma_{t+1} = \alpha y_t + (1-\alpha) \gamma_t$ que exige disponer de un valor de α y de un valor inicial de predicción γ_t . Esta última cuestión es relativamente poco importante y puede resolverse adoptando directamente $\gamma_1 = y_1$. En cuanto al valor de α , se calcula el valor que minimiza el error cuadrático medio.

ESCENARIO: *Predicción usando Alisado Exponencial*

1. El usuario busca la *serie de datos* con la cual trabajar
2. *Graficar* la serie.
3. Crear el *modelo* de alisado exponencial y calcular el valor óptimo de α .
4. Realizar la predicción

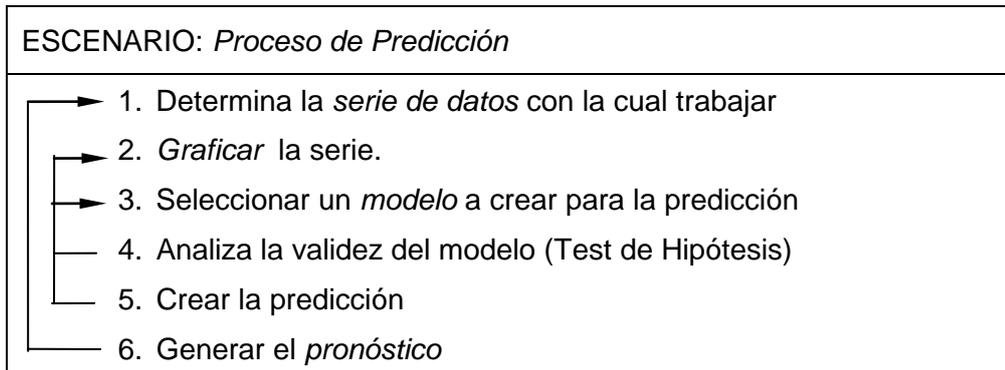
En la Figura 2 se muestra el *Diagrama de Objetos* correspondiente, el cual representa la interacción entre los objetos que intervienen en el escenarios.



Presentamos el escenario anterior como ejemplo, no obstante, para desarrollar este trabajo hemos analizado los escenarios de predicción asociados a los siguientes casos:

- Predicción aplicando el modelo de Medias Móviles y el modelo de Alisado Exponencial a una serie de datos almacenados en una Base de Datos.
- Aplicación del modelo de Ajuste de tendencia para determinar la tendencia a largo plazo de la demanda de un producto.
- Aplicación de la técnica Delfos para asignar valor a una variable exógena de un modelo econométrico (Pindyck R.y otro,1980) de predicción a mediano plazo.
- Aplicación de modelos Arimas para la predicción a partir de una serie de datos.
- Uso de modelos uniecuacionales para relacionar un forma lineal una variable endógena y diversas variables exógenas.

Como resultado del análisis de estos casos, hemos diseñado el siguiente escenario principal correspondiente a un *proceso de predicción*:



De este escenario, un análisis de sustantivos, permite inferir como clases/objetos candidatos los siguientes: *Pronósticos*, *Modelos*, *Series*, *Elementos de Serie*, *Gráficos* y *Test de Hipótesis*. En la Figura 3 se muestra el *Diagrama de Interacción* correspondiente, el cual representa la interacción entre las clases/objetos candidatos.

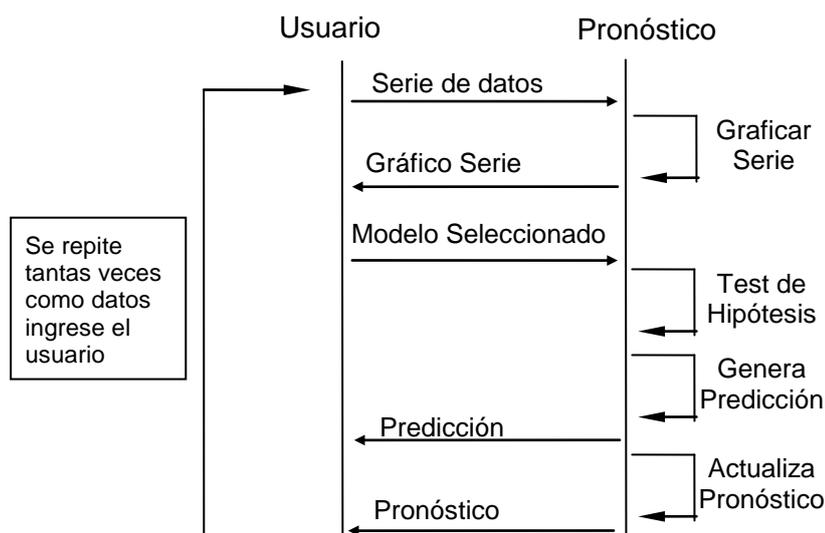


Figura 4: Diagrama de Interacción

En este trabajo definimos como *pronóstico* al resultado de un proceso de predicción. Este resultado por cierto no es único ni permanente, sino que generalmente consiste en un conjunto de alternativas, cada una de las cuales depende del entorno en el cual se enmarca la predicción y de la técnica utilizada. Luego, la clase *Pronóstico* inferida a partir del escenario anterior, consistirá en un conjunto de n predicciones. En particular, un pronóstico podrá estar asociado a n predicciones. De este modo inferimos que existe una relación de asociación entre la clase *Pronóstico* y la clase *Predicciones*.

Una predicción es en si una *serie* de datos, por lo tanto toda predicción estará asociada a una serie de datos. Inferimos entonces una relación de asociación entre las clases *Predicciones* y *Series*. Es también de considerar que toda predicción resulta de aplicar algún modelo, por lo tanto cada una de ellas está asociada a un modelo especialmente creado y validado. Luego la clase *Predicciones* se relaciona a la clase *Modelos* mediante una relación de asociación.

Por otra parte, la validación de un modelo como condición para aceptar la predicción realizada por el mismo, implica que debe satisfacer los test de hipótesis asociados a dicho modelo. De este modo, se infiere también una relación de asociación entre las clase *Modelos* y la clase *Test de Hipótesis*.

Para realizar la predicción, el modelo seleccionado, hace uso de la serie que contiene los datos base de la predicción. De este modo, se infiere una relación de uso entre la clase *Modelos* y la clase *Series*.

Consideramos a una serie como una estructura con la cual se representa un conjunto ordenado de datos. Cada uno de estos datos constituye un elemento de la serie. En base a esta definición, se establece una relación de composición entre la clase *Series* y la clase *Elementos de Serie*.

Una serie puede mostrarse al usuario en forma de gráfico. De este modo, cada serie tendrá un gráfico asociado. Por lo tanto inferimos una relación de asociación entre las clase *Series* y la clase *Gráficos*.

Representamos en el Diagrama de Clases 1 los resultados de este análisis.

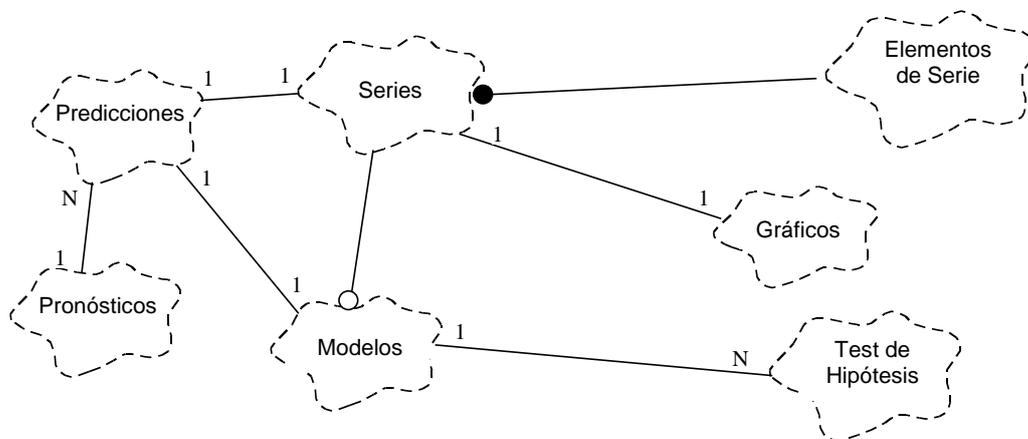


Diagrama de Clases 1

En todo proceso de predicción, es necesario realizar un análisis de la validez del modelo utilizado a partir de ciertas medidas de error. En otros términos, la aplicación de un modelo de predicciones implícitamente supone que el mismo debe ser sometido a ciertos test de hipótesis. Las características de cada modelo en general determinan cuales son los análisis de validación aconsejables, no obstante es el experto quien decide finalmente el conjunto de test de hipótesis que debe verificar el modelo. En el diseño de este sistema hemos definido a los métodos de validación como una clase separada del modelo, debido a que son mucho los test a utilizar, y todos ellos comparten atributos y métodos. Por otra parte, se tiene una arquitectura más flexible. En esta versión prototipo del SSP hemos incluido los siguientes test:

- *Test de la Media*: Test de hipótesis basado en el cálculo de la media estadística de una serie.
- *Test de la Varianza*: Test de hipótesis basado en el cálculo de la varianza de una serie, que utiliza la tabla de chi_cuadrado.
- *Test de Correlación*: Test de hipótesis basado en el cálculo de la correlación estadística de una serie.
- *Contraste T*: Test de hipótesis basado en la tabla de t-Student
- *Contraste F*: Test de hipótesis para validar los parámetros de los modelos.

Estos tests comparten atributos y métodos, no obstante poseen otros que son propios de cada uno, por esta razón hemos definido las clases: *Test Media*, *Test Varianza*, *Test Correlación*, *Contraste T* y *Contraste F*. Cada una de estas clases hereda de la clase *Test de Hipótesis* los

atributos y métodos comunes. Existe por lo tanto una relación de herencia. Mostramos esto en el Diagrama de Clases 2:

En el escenario anterior hemos considerado a una serie como una estructura que representa un

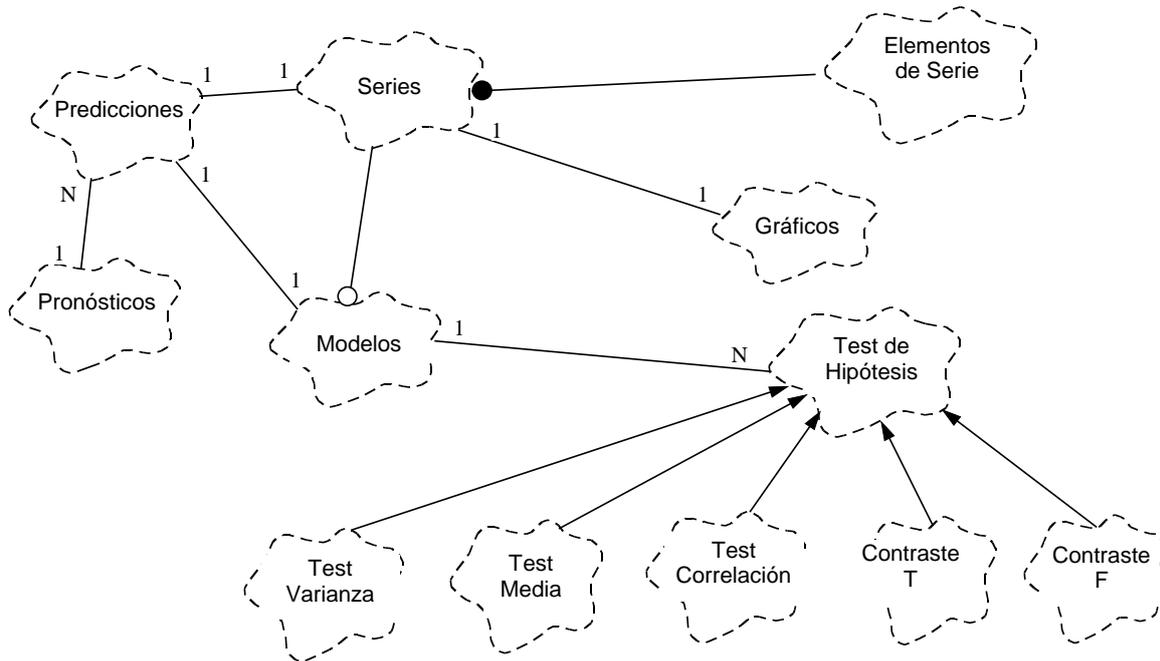


Diagrama de Clases 2

conjunto ordenado de datos sin analizar el tipo de datos. En un procesos de predicción se utilizan dos tipos de series de datos:

- Las *Series Estadísticas*, conformadas por los valores de una variable con sus respectivos datos estadísticos para ser usada en un análisis de tipo estadístico. Este tipo de series es generada por las técnicas de predicción en base a información subjetiva, como el caso de la técnica Delfos.
- Las *Series Temporales*, conformadas por los valores de una variable en función del tiempo con sus estadísticos, para ser usada a efectos de predicción con algún modelo matemático de predicción.

Si bien estas series comparten la estructura, poseen diferencias sustanciales en cuanto a atributos y métodos asociados. Por esta razón hemos definido las clases *Series Estadísticas* y *Series Temporales*. Estas clases se vinculan con la clase *Series* a través de una relación de herencia. Representamos esto en el Diagrama de Clases 3:

Cabe destacar que la clase *Modelos* inferida del escenario anterior, es una abstracción de todas las técnicas, métodos y modelos que el experto puede utilizar para generar un pronóstico a partir de información subjetiva, histórica o mediante modelos de relaciones causales. Hemos definido una clase para cada tipo de técnica y modelo, a su vez, hemos agrupado a través de clases tipos de modelos que comparten atributos y estructuras. De este modo, a partir de la clase *Modelos* hemos generado la jerarquía de clases, vinculadas mediante relación de herencia, que presenta el Diagrama de Clases 4:

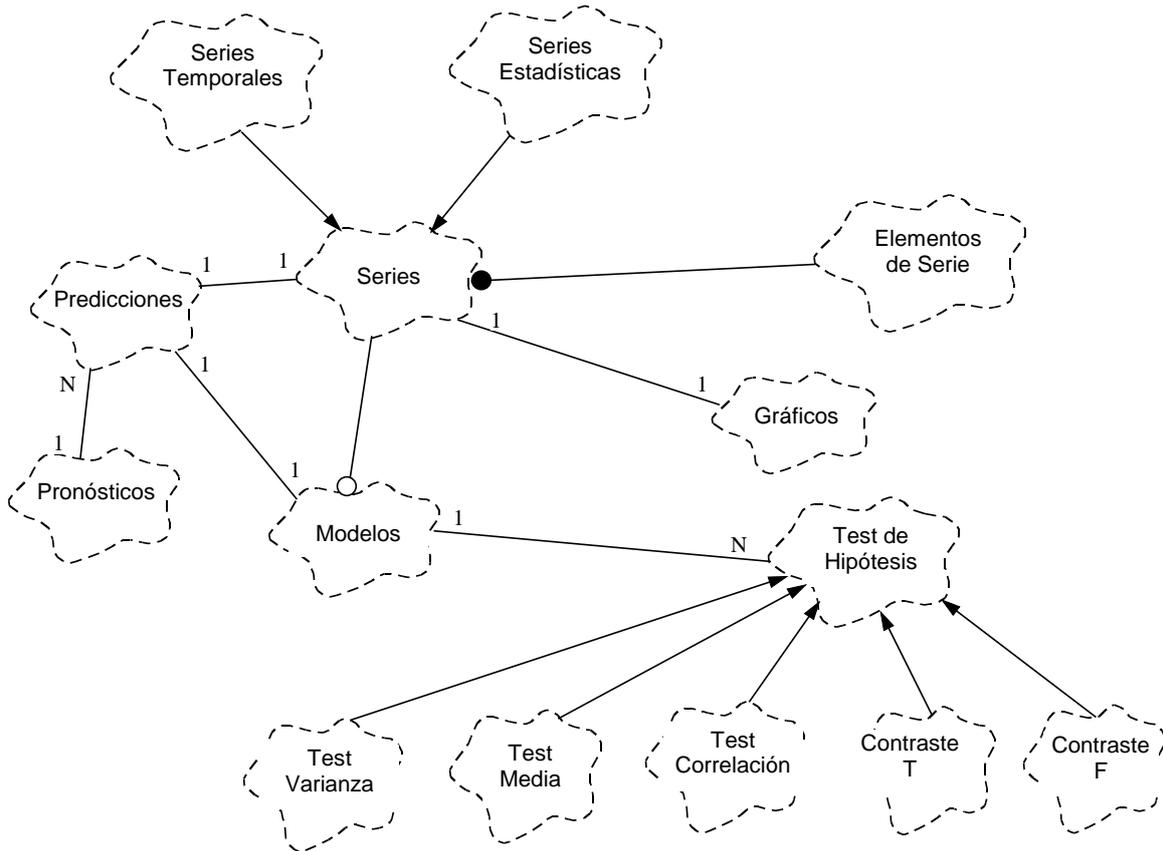


Diagrama de Clases 3

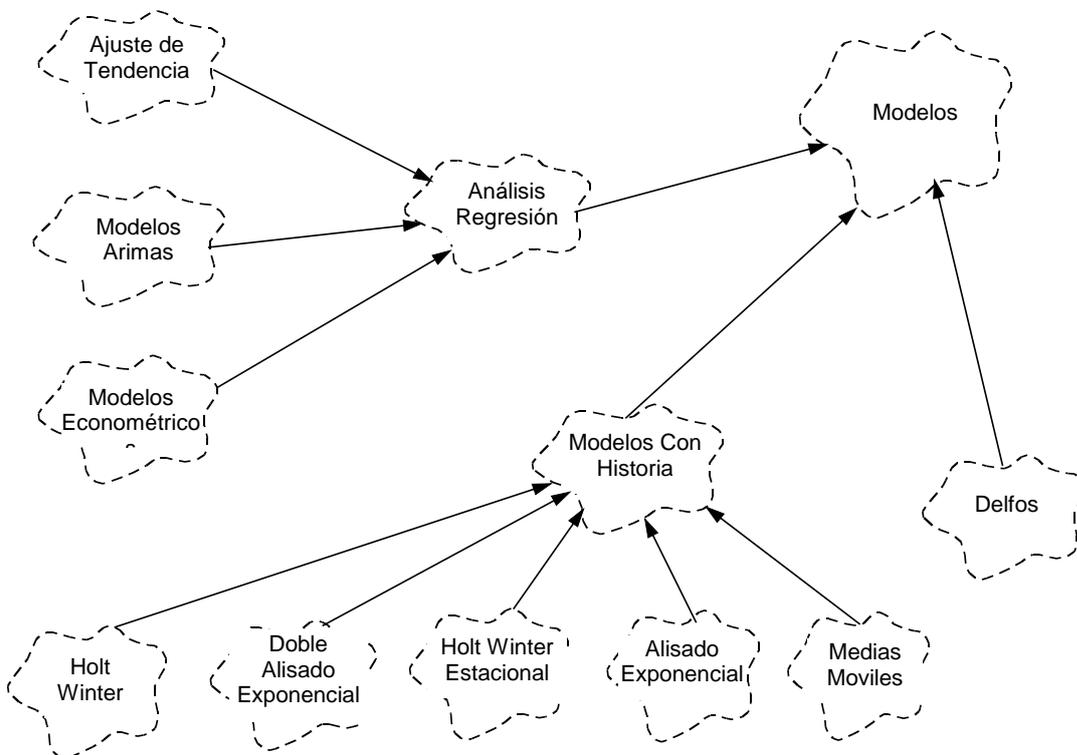


Diagrama de Clases 4

Hemos presentado así una comprensión básica del problema. El paso siguiente fue realizar un análisis detallado de la estructura de cada clase y de sus responsabilidades (métodos). Para esto

hemos desarrollado el *Diccionario de Clase* para cada una de las clases. A modo de ejemplo se muestra el diccionario de la clase *Series Temporales*.

NOMBRE DE CLASE: <i>Series Temporales</i>
DESCRIPCION: <i>Contiene los valores de una variable en función del tiempo con sus estadísticos para ser usada a efectos de predicción con algún modelo matemático de predicción.</i>
ATRIBUTOS: <ul style="list-style-type: none">➤ <i>Lista de Elementos Temporarios</i>➤ <i>Tendencia</i>➤ <i>Estacionalidad</i>➤ <i>Ecuación</i>
METODOS: <ul style="list-style-type: none">➤ <i>Crear Serie</i>➤ <i>ABM Valores</i>➤ <i>Calcular Estadísticos</i>➤ <i>Interpolar</i>➤ <i>Calcular Medias Móviles (Nro_Orden)</i>➤ <i>Diferencias Sucesivas</i>➤ <i>Transformación Logarítmica</i>➤ <i>Calcular_Pendiente</i>➤ <i>Calcular_Estacionalidad</i>➤ <i>F_Suma</i>➤ <i>F_Resta</i>➤ <i>F_Producto</i>➤ <i>F_División</i>➤ <i>F_Potencia</i>➤ <i>F_Raíz</i>
RELACIONES DE HERENCIA: SUPERCLASE: <i>Series</i>

A partir de este análisis detallado se obtuvo información sobre las relaciones estructurales necesarias (por ejemplo, que la clase *Series Temporales* necesitará de una lista de los elementos temporales que la conforma), como así también un detalle de los atributos de cada clase, y un detalle de los métodos requeridos por la clase.

CONCLUSIONES

Como conclusión del presente trabajo debemos destacar en primer lugar que, conforme a lo propuesto por Sage 1991, hemos adoptado una política de desarrollo evolutiva del SSP prototipo. Por esto, el análisis y diseño del subsistema fue avanzado hasta un nivel considerado satisfactorio y luego pasamos a la etapa de programación. Los resultados de esta primer versión académica del prototipo posiblemente sugieran cambios tanto en la estructura de la información como del diseño del sistema. En este trabajo hemos realizado una breve descripción de las características de un proceso de predicción, y en función de las mismas hemos descripto los pasos seguidos y los criterios de análisis utilizados.

Consideramos que la arquitectura definida, en la cual hemos separado en clases a los modelos, métodos y técnicas de predicción, a los métodos de validación y a los datos a través de las series, ha permitido lograr una arquitectura muy flexible, ya que no está vinculada a un tipo de predicción,

a un modelo de datos, ni a un tipo de pronóstico. Esta estructura permite además que el experto pueda incorporar sus propios criterios de valoraciones en las distintas etapas de un proceso de predicciones, como así también establecer el diseño de pronóstico que considere conveniente.

También es importante destacar que el diseño del módulo compatible con la estructura de un SSD global, establece un vínculo muy fluido de comunicación con los módulos a los cuales debe suministrar información. Este vínculo es el que se requiere para trabajar con predicciones sometidas a procesos de actualización continua, necesarios en los entornos cambiantes en los que se basan las predicciones.

El prototipo se implementa utilizando el lenguaje de programación visual DELPHI, para lo cual debimos compatibilizar el diseño de objetos realizado siguiendo la metodología de G. Booch, y el lenguaje de programación. La adaptación más importante ha sido quizás la necesidad de almacenar objetos en tablas relacionales.

REFERENCIAS

- Adamopoulos G., N. Karacapilidis & S. Pantazopoulos, (1994), *Production Management in the Textil Industry Using the "YFADI" Decision Support System*, Comp. & Chem. Engng Vol **18**, pp s577-s583.
- Booch,G., (1993), *Object Oriented Analysis and Design*, J.Wiley & Sons
- Lo,T. (1994), *An Expert Ssystem for Choosing Demand Forecasting Techniques*, Intern. Journal of Production Economics. pp.5-16.
- Pindyck,R. y Rubinfeld,D.,(1980), *Modelos Económicos*, Labor S.A., Madrid.
- Pulido, A., (1989), *Predicción Económicas y Empresariales*. Pirámide S.A., Madrid
- Rico M., O.Yuschak, M.L.Taverna, J. C.Ramos, M.R Galli and O.Chiotti, *Decision Support Systems Generator for Industrial Companies*, Proceeding of 21th International Sage,A.P. , *Decision Support Systems Engineering*, John Wiley & Sons (1991).