

ENTREVISTA A ALBERTO ESPINOSA

Biografía: El profesor Espinosa tiene un doctorado y una maestría en Sistemas de Información por Carnegie Mellon University, una maestría en Administración por Texas Tech University y un título en Ingeniería Mecánica por la Pontificia Universidad Católica del Perú. Diseñó el programa de Maestría en Analítica y la especialización en Analítica de Negocios en Kogod, además de programas en Tecnología de la Información y Analítica. Con más de 70 publicaciones, ha coescrito libros sobre coordinación en zonas horarias, big data y machine learning en R. Su investigación aborda la coordinación en proyectos técnicos globales y el análisis de redes sociales y grafos. Ha publicado en revistas líderes como Management Science y Information Systems Research, y ha presentado en conferencias de prestigio. Enseña analítica predictiva, machine learning, redes organizacionales y programación en R. Cuenta con experiencia como ingeniero de diseño y en altos cargos directivos en organizaciones internacionales, supervisando TI global, gestión de datos y desarrollo de software para contabilidad y finanzas.

Rafael Andres Cangahuala Martinez

20207560 | a20207560@pucp.edu.pe

Estudiante del séptimo ciclo de la Facultad de Ciencias Contables

Miembro del equipo de Costos

Anggi Janet Leiva Saavedra

20222763 | a20222763@pucp.edu.pe

Estudiante del quinto ciclo de la Facultad de Ciencias Contables

Miembro del equipo de Finanzas

¿Cuáles son las principales ventajas de aplicar técnica de análisis predictivo y aprendizaje automático en la toma de decisiones de los managers?

Hoy en día, el uso de datos no es solo una ventaja, sino una necesidad. Con la cantidad masiva de datos que se recolectan en diversas áreas, como ventas online, datos médicos y relacionados con pandemias, es esencial aprender a navegar y trabajar con esta información. En este contexto, la analítica predictiva juega un papel crucial. Esta técnica, que emplea *machine learning*, permite

analizar datos y predecir resultados de interés, como, por ejemplo, prever ventas que han desaparecido, o determinar la probabilidad de que un paciente con COVID termine siendo hospitalizado o necesite atención en emergencias. También puede predecir la probabilidad de que un candidato gane una elección, o si una persona con un diagnóstico de tumor positivo desarrollará complicaciones.

Aunque las predicciones y análisis de este tipo no son nuevos y se han hecho durante años utilizando estadísticas tradicionales y modelos de regresión, el uso de

machine learning ha mejorado notablemente la precisión de estas predicciones. Por ejemplo, un médico que examina una radiografía puede detectar un tumor comparando la imagen con muchas otras que ha visto a lo largo de su carrera. Sin embargo, un sistema basado en *machine learning* puede comparar una radiografía con millones de imágenes digitalizadas, haciendo el proceso mucho más exacto y exhaustivo. Esto no reemplaza al médico, pero sí lo ayuda a tomar decisiones más informadas.

Hoy en día, prácticamente todo se basa en datos y análisis. En Estados Unidos, por ejemplo, no puedes sentarte frente a un cliente o gerente y decir "yo pienso que esto es así porque tengo experiencia". Las personas quieren ver los datos que respaldan las decisiones. La analítica predictiva, por lo tanto, está presente en múltiples áreas: desde la medicina y el marketing hasta las finanzas y la contabilidad forense, especialmente en la detección de fraudes.

Nuestro programa de *Predictive Analytics* está orientado a gerentes y profesionales analíticos, no a científicos de datos que se dedican a las complejas ecuaciones. Está diseñado para aquellos que toman decisiones en las empresas. Esta tendencia es algo que, con el tiempo, se adaptará a todos los tipos de negocios e industrias. Es crucial que los profesionales, desde ya, adquieran conocimientos en bases de datos y software, ya que el *machine learning* está convirtiéndose en el futuro de todo el sector corporativo.

¿Cómo puede el aprendizaje automático mejorar la eficacia de las auditorías y detectar el fraude?

El aprendizaje automático (*machine learning*) es especialmente útil en la detección

de fraudes. La clave es que los modelos utilizados para la detección de fraudes son muy similares a los de predicción. La principal ventaja es la capacidad de comparar rápidamente grandes volúmenes de datos históricos con nuevos casos. Por ejemplo, en el caso de fraude, el sistema puede detectar un patrón típico de transacción fraudulenta al compararlo con millones de observaciones anteriores. Este proceso no solo es útil en fraudes financieros, sino también en otros tipos de fraude, como el fraude contable, o incluso en la detección de plagio en las clases.

Lo que hacen los programas de *machine learning* es comparar patrones y detectar la probabilidad de que una transacción o situación sea fraudulenta. En resumen, el aprendizaje automático agiliza este proceso de detección, no solo porque mejora la velocidad, sino también la precisión en comparación con los métodos tradicionales. Durante años, la estadística y los modelos de regresión se han utilizado para este tipo de predicciones. Por ejemplo, con 100 000 observaciones y 200 variables en la base de datos, se usaba un modelo de regresión para predecir un resultado, como detectar si una transacción es fraudulenta o no.

En los modelos tradicionales, simplemente analizábamos la relación entre las variables predictoras y el resultado, basándonos en datos históricos. Sin embargo, el aprendizaje automático introduce un enfoque diferente. En *machine learning*, los datos se dividen en dos partes: una para entrenar el modelo (como en la estadística) y otra parte que se usa para probarlo. Luego, el modelo entrenado se aplica a los datos no utilizados y se evalúa su capacidad de predicción. Esto permite evaluar qué tan preciso es el modelo cuando predice con datos que no se utilizaron en su construcción.

Este enfoque es mucho más exacto cuando se enfrentan datos nuevos, ya que el modelo se evalúa continuamente y se ajusta. Además, la programación en *machine learning* es muy fácil de automatizar. Por ejemplo, se puede dividir una base de datos en un 70% para entrenar el modelo y un 30% para probarlo. Este proceso se repite miles de veces, y se calcula el error promedio para obtener un modelo más preciso. Este procedimiento se puede automatizar, lo que permite que el sistema realice el análisis de manera independiente sin necesidad de intervención manual.

Aparte de esto, herramientas automáticas como *AutoML* facilitan aún más el proceso. *AutoML* es un programa que toma una base de datos y automáticamente prueba decenas de modelos diferentes. Luego, evalúa los resultados utilizando técnicas como la validación cruzada (*cross-validation*). El usuario solo tiene que esperar a que el programa termine para obtener un modelo perfecto para sus necesidades. Esta automatización y la creación de modelos sofisticados se han vuelto posibles gracias al avance de *machine learning*.

¿Qué tipo de decisiones estratégicas se pueden mejorar con el uso de análisis predictivos?

El modelo predictivo es parte fundamental de lo que conocemos como *analytics* o análisis de datos. Este proceso implica extraer el significado de los datos para tomar decisiones informadas. Por ejemplo, si te doy una tabla con 10 millones de observaciones, el primer reto es entender qué significan esos datos. Los números por sí mismos no tienen emoción ni contexto, son fríos y abstractos. Sin embargo, al aplicar *machine learning* a esos datos y construir un modelo de analíti-

ca predictiva, podemos generar predicciones, evaluar resultados y analizar qué variables predicen otras.

De esta manera, podemos construir lo que se conoce como “storytelling” o narrativas, que nos permiten entender qué factores están afectando las ventas, el crédito, la propagación de una pandemia o incluso la malignidad de un tumor. A través de esta interpretación de datos, los analistas pueden presentar las implicaciones de los resultados de manera comprensible, ya sea en inglés o en castellano, para que un gerente o tomador de decisiones pueda entender y actuar en consecuencia.

Como analista, es necesario tener una base técnica sólida que incluya conocimientos de estadística, *machine learning* y programación. Sin embargo, el objetivo no es que el *machine learning* desarrolle estrategias por sí solo, sino que informe al gerente para que este pueda tomar decisiones estratégicas bien fundamentadas. El modelo predictivo ayuda a proporcionar insights y análisis, pero, al final, es el capital humano quien decide cómo utilizar esa información para generar acciones concretas.

Es interesante observar cómo, aunque el *machine learning* tiene la capacidad de procesar datos a gran escala y generar conclusiones significativas, la toma de decisiones sigue siendo una tarea del ser humano. Sin embargo, no todas las empresas cuentan con estas herramientas avanzadas, y aunque el desarrollo de estas tecnologías es continuo, la implementación todavía es un desafío para muchas organizaciones.

¿Qué consejo le darías a un manager que desea implementar análisis predictivo y aprendizaje automático en su empresa, pero no sabe por dónde empezar?

En principio, tengo una pequeña discrepancia con lo que mencionaste, ya que hoy en día la situación es todo lo contrario. Aunque el *machine learning*, la inteligencia artificial y la analítica predictiva han existido durante muchos años, anteriormente se necesitaban recursos de cómputo complejos, como computadoras mainframe, y contar con personal altamente capacitado para programar y gestionar estos sistemas.

Esto me recuerda al teléfono en sus primeros días: una tecnología extremadamente compleja que requería operadores para conectar las llamadas. Hoy en día, sin embargo, la tecnología detrás del teléfono es igualmente sofisticada, pero lo que el usuario final experimenta es simple y accesible: basta con marcar un número y realizar una llamada. Lo mismo está ocurriendo con *machine learning* y la inteligencia artificial. Las herramientas son cada vez más fáciles de usar, lo que las hace más accesibles.

Hoy en día, la inteligencia artificial no reemplaza a los seres humanos, sino que quienes dominarán el futuro serán aquellos que sepan cómo utilizar estas herramientas. Para los estudiantes y profesionales, esto significa que aprender sobre estas tecnologías es fundamental, ya que son las herramientas del futuro. Es similar a ser un carpintero que usaba martillos y alicates durante años, pero que ahora tiene acceso a herramientas eléctricas como taladros neumáticos y cepilladoras eléctricas. Estas herramientas hacen el trabajo mucho más rápido y eficiente. Cualquier persona con un conocimiento básico de carpintería, tras un entrenamiento adecuado, puede aprender a usar estas máquinas y producir muebles en un tiempo mucho más corto que antes, cuando todo se hacía manualmente.

Lo mismo está sucediendo con el *machine learning* y la inteligencia artificial: las herramientas son más poderosas, potentes y fáciles de usar, lo que facilita mucho el trabajo. Sin embargo, esto también trae consigo nuevos retos, los cuales queremos explorar, particularmente, en el contexto de la contabilidad.

En el contexto de la contabilidad, ¿cuáles son los principales riesgos de confiar exclusivamente en análisis predictivo y cómo pueden mitigarse?

Considero que el principal riesgo asociado al uso de la inteligencia artificial y el *machine learning* es el factor humano. Un ejemplo claro de esto es dejar que la inteligencia artificial responda preguntas cruciales como “¿dónde debo invertir en el mercado de valores?” o “¿qué decisión debo tomar?”.

Recuerdo que hace aproximadamente 30 años, cuando estaba haciendo mi maestría, uno de mis profesores nos dejó una lección fundamental. Al comenzar la clase, escribió en la pizarra “MD al cubo” y me preguntó qué significaba. Al principio no lo entendí, pero luego explicó: “MD significa *Models don't make decisions, managers do*” (Los modelos no toman decisiones, los gerentes sí). Lo que quiero decir con esto es que los modelos están diseñados para apoyar la toma de decisiones, pero la decisión final debe ser humana.

Para mí, el verdadero riesgo es permitir que los modelos tomen decisiones, especialmente en situaciones complejas. Existen decisiones simples, como la recomendación de productos en Amazon basada en patrones de compra anteriores, donde la automatización puede ser útil. Sin embargo, en situaciones más delicadas, como un fraude

o la decisión de demandar a alguien, no podemos dejar que el modelo decida; debe ser el ser humano quien tome esa decisión.

En una conferencia, un científico de datos presentó un estudio en el que su principal preocupación era la precisión de las predicciones, sin importar tanto la interpretación de los resultados. A continuación, se presentó un equipo de trabajo del banco Capital One que tenía una perspectiva diferente. Ellos explicaron que, cuando rechazan un préstamo, deben ser capaces de explicar el porqué tanto a los reguladores como al cliente, y aquí la intervención humana es crucial. Si se utiliza un modelo automatizado para decidir quién recibe un préstamo, puede darse el caso de discriminación o de tomar decisiones erróneas, como negar un préstamo a alguien con buen crédito debido a patrones sesgados. Por eso, la intervención humana es esencial.

Por supuesto, las decisiones pequeñas, como las recomendaciones de productos, pueden ser automatizadas, pero las decisiones estratégicas y complejas siempre deben involucrar un factor humano.

Un colega mío, que ocupa el puesto de Innovation Officer en una startup, y que ha trabajado con varias consultoras de grandes empresas, compartió una reflexión interesante. Comentó que, aunque utilizan *machine learning* para crear modelos que detecten fraudes y realicen análisis de riesgo para préstamos, reconocen que el modelo solo debe realizar las tareas automáticas y mecánicas. La intervención humana es necesaria para interpretar los resultados, tomar decisiones informadas y explicar los resultados a las partes involucradas.

Me parece fascinante cómo, a pesar del rápido desarrollo de las tecnologías, siem-

pre habrá un factor humano que juega un papel fundamental en la toma de decisiones, especialmente, cuando trabajamos con personas a las que ofrecemos servicios. Ese factor humano es lo que, al final, marca la diferencia en la decisión final.

Aunque las tecnologías avanzan y tareas más sencillas se pueden automatizar, las decisiones importantes siempre dependerán del juicio humano. El principal riesgo, entonces, es permitir que los modelos tomen decisiones directamente, sin que los gerentes intervengan. Los modelos deben ser una herramienta para apoyar la toma de decisiones, no una solución automática que reemplace la intervención humana.

Por ejemplo, en mis clases, cuando propongo problemas o casos prácticos, muchas veces los estudiantes sugieren soluciones automáticas, como sistemas que hagan algo de forma automática. Sin embargo, siempre les recuerdo que, en el mundo real, los gerentes prefieren tomar decisiones manualmente. Les doy un ejemplo: desarrollé un programa para una agencia que otorgaba ayudas para calefacción a personas necesitadas. El sistema tomaba los datos de los solicitantes y evaluaba si calificaban o no para la ayuda. Sin embargo, cuando el sistema recomendaba negar la ayuda a alguien, era crucial que un ser humano revisara el caso, ya que esa decisión podía afectar directamente la vida de la persona. Si una persona necesitaba calefacción, no se podía permitir que un sistema automatizado tomara la decisión sin una intervención humana que garantizara la corrección de la evaluación.

Los datos y los análisis son herramientas valiosas, pero no deben reemplazar el juicio humano en situaciones críticas.