



# **Sistema de Control de Calidad para Cervecería**

## **Proyecto Final** Datos Masivos

Por:

Guzmán Domínguez Andrés Manuel

Berzunza Rodríguez Alondra Vanianinetl

22 de mayo de 2025

El proyecto Sistema de Control de Calidad para Cervecería propone un monitoreo inteligente en tiempo real para optimizar la producción cervecera mediante tecnología de vanguardia. Este sistema integral combina sensores IoT, inteligencia artificial y principios de SCADA para supervisar y predecir la calidad de cada lote de cerveza durante su elaboración. Con esta solución, se busca mejorar la consistencia del producto final, reducir pérdidas por lotes defectuosos y habilitar ajustes proactivos en el proceso de fermentación, todo sustentado por una arquitectura tecnológica moderna y robusta.

## Marco Teórico

En el Internet de las Cosas (IoT), dispositivos y sensores conectados recopilan datos del entorno físico en tiempo real. En una cervecería industrial, esto implica sensores de temperatura, pH, nivel de CO<sub>2</sub>, presión, entre otros, instalados en tanques de fermentación y líneas de producción. Estos sensores IoT transmiten constantes flujos de datos que, integrados a un sistema SCADA (Supervisory Control and Data Acquisition), permiten la supervisión y control continuo del proceso.

Un SCADA típico en cervecería monitorea variables críticas y envía alertas al operador o activa actuadores (válvulas, enfriadores) cuando las lecturas salen de rangos óptimos. La combinación de IoT y SCADA proporciona la infraestructura base para capturar información detallada del proceso cervecero en tiempo real, sirviendo como columna vertebral de un sistema de control de calidad automatizado. De hecho, el sistema desarrollado se conecta directamente con los sensores de la línea de producción y se sincroniza con sistemas MES (Manufacturing Execution System) existentes, alimentando una base de datos central con todos los parámetros recolectados.

La Inteligencia Artificial (IA) aplicada a la predicción de calidad utiliza algoritmos de machine learning que analizan datos históricos y en vivo para encontrar patrones complejos que los métodos tradicionales no detectan. En el contexto cervecero, modelos de IA pueden correlacionar las condiciones de fermentación y las propiedades de los ingredientes con métricas finales de calidad (sabor, grado alcohólico, amargor, color, etc.). Mediante técnicas de aprendizaje automático, el sistema puede predecir la calidad de un lote antes de concluir su producción, anticipándose a posibles desviaciones. Esto entra dentro de la tendencia de industria 4.0, donde la analítica predictiva optimiza procesos industriales. Al entrenar modelos con grandes volúmenes de datos históricos (miles o millones de registros de lotes anteriores), la IA es capaz de generalizar y estimar con alta precisión resultados como el puntaje de calidad de la cerveza.

Estas predicciones permiten implementar un control de calidad preventivo: en lugar de detectar problemas al final (cuando el producto ya está terminado y podría desecharse), se detectan patrones anómalos durante el proceso, habilitando correcciones en tiempo real. El Control de Calidad en procesos cerveceros tradicionalmente combina mediciones de laboratorio (densidad,

pH, IBU de amargor, contenido alcohólico) con catas sensoriales. Sin embargo, estos métodos pueden ser lentos y reactivos. Un enfoque automatizado y en tiempo real aprovecha los datos de sensores industriales para asegurar que cada etapa del proceso (maceración, fermentación, filtrado, envasado) se mantenga dentro de especificaciones óptimas. Variables como la temperatura de fermentación, el tiempo de fermentación, las proporciones de malta, lúpulo y levadura, y las pérdidas en cada etapa, son indicadores directos de la calidad final. Si estas variables se controlan estrechamente (mediante IoT/SCADA) y sus datos se analizan con IA, es posible garantizar una mayor consistencia lote a lote. En suma, el marco teórico del proyecto se apoya en la convergencia de IoT (para recolección de datos en planta), SCADA (para supervisión y control en tiempo real) e Inteligencia Artificial (para análisis predictivo), todo orientado a fortalecer el aseguramiento de la calidad en la producción cervecera.

## Planteamiento del Problema

En la producción cervecera actual se identificaron varios problemas que motivan este proyecto:

- **Control de calidad impreciso:** Los métodos tradicionales no garantizan la detección oportuna de todas las variaciones en el proceso, lo que conduce a inconsistencias en el producto final.
- **Detección tardía de anomalías:** Muchas irregularidades durante la fermentación u otras etapas se descubren demasiado tarde, cuando el lote ya está fuera de especificación.
- **Pérdidas económicas por lotes rechazados:** La falta de detección temprana conlleva que lotes completos deban ser descartados por baja calidad, generando desperdicio de insumos y pérdidas financieras.

Estos problemas afectan tanto la eficiencia como la rentabilidad de la cervecería. Un control de calidad manual o reactivo implica que se toman acciones correctivas solo después de que el daño está hecho (por ejemplo, cuando un lote ya no cumple estándares y debe eliminarse). La imprecisión y demora en la retroalimentación dificultan optimizar el proceso en tiempo real. Por ello, surge la necesidad de una solución tecnológica más avanzada que permita monitoreo continuo y predictivo, capaz de anticipar las desviaciones antes de que se conviertan en fallas de calidad.

## **Hipótesis**

Se plantea la hipótesis de que integrar un sistema de inteligencia artificial con monitoreo en tiempo real mejorará significativamente el control de calidad en la producción cervecera. Es decir, si se implementa un sistema automatizado con IA para la predicción y monitoreo de la calidad, entonces se podrá detectar tempranamente cualquier variación anómala en parámetros críticos del proceso y ajustar las condiciones de producción a tiempo. Como resultado, disminuirán los lotes rechazados y las pérdidas asociadas, a la vez que aumentará la consistencia y calidad promedio de la cerveza producida.

En resumen, la hipótesis anticipa que una solución de IoT + IA aplicada al control cervecero optimizará los resultados de producción, resolviendo las deficiencias del método tradicional.

## Desarrollo

El desarrollo del proyecto abarcó el diseño e implementación de una arquitectura técnica conformada por varios componentes integrados. En esencia, el sistema se compone de tres módulos principales:

- **Pipeline de Datos y Preprocesamiento:** Encargado de captar los datos desde la planta (por ejemplo, lecturas de sensores IoT integrados al SCADA) y realizar la depuración inicial. Incluye pasos de limpieza (p. ej., eliminación de valores atípicos o erróneos), transformación de datos (normalización de escalas, conversión de formatos de fecha, separación de campos compuestos) y validación de la integridad de la información. Este pipeline se implementó sobre Apache Spark, lo que permite procesar grandes volúmenes de datos de producción de forma distribuida y eficiente. Los datos históricos de producción (por ejemplo, registros de más de 1 millón de lotes) fueron almacenados en formato Parquet para facilitar su lectura y manipulación a gran escala. Cada nuevo lote generado en la planta alimenta este pipeline en tiempo real, garantizando que la información esté siempre actualizada para el modelo de IA.
- **Modelo de Predicción de Calidad:** Es el núcleo de inteligencia artificial del sistema. Se entrenó un modelo de aprendizaje automático capaz de predecir la calidad de la cerveza (cuantificada mediante un puntaje numérico) a partir de las variables del proceso. En concreto, se utilizó un algoritmo de árboles potenciados en gradiente (Gradient Boosted Trees) configurado como regresor. Este modelo toma como features entradas todas las mediciones relevantes de cada lote (tiempo y temperatura de fermentación, pH, densidad original, proporción de ingredientes malta/lúpulo/levadura, pérdidas en proceso, estilo de cerveza, etc.) y estima un `Quality_Score` para el lote. Para entrenarlo adecuadamente, se dividieron los datos históricos en conjuntos de entrenamiento, validación y prueba de forma temporal (usando datos de años anteriores para entrenar, y los más recientes para validar).

Además, se aplicaron técnicas de balanceo de datos para que el modelo no estuviera sesgado: por ejemplo, se sub-muestrearon los casos de calidad alta predominantes para equilibrarlos con casos de calidad baja más escasos. También se llevó a cabo fine-tuning de hiperparámetros del modelo para optimizar su desempeño predictivo. El modelo final entrenado alcanzó altos

niveles de precisión (como se detalla en Resultados) y fue almacenado en disco para su uso en la fase de inferencia en tiempo real.

- **Panel de Control en Tiempo Real:** Es la interfaz de visualización y monitoreo del sistema, desarrollada como una aplicación web usando Streamlit. El panel emula las funciones de un SCADA moderno, presentando en pantalla las variables críticas del proceso cervecero y las predicciones de calidad de cada lote en producción, todo actualizado en tiempo real. La interfaz está diseñada de forma intuitiva, permitiendo al operador de planta ver tendencias, gráficas y alertas de un vistazo. Por ejemplo, se muestran indicadores de la puntuación de calidad por lote con códigos de color (verde óptimo, amarillo aceptable, rojo crítico) y se activa un sistema de alertas tempranas en caso de que la predicción de calidad caiga por debajo de un umbral definido. El panel también opera 24/7 sin interrupciones, lo que significa que el monitoreo es continuo incluso en turnos nocturnos o fines de semana. Para simular la recepción continua de datos, se implementó un mecanismo de micro-lotes de datos: el sistema toma periódicamente pequeñas muestras del dataset de producción (por ejemplo, cada 2 segundos) y las procesa como si fuesen lecturas en vivo, actualizando las predicciones en el dashboard.

Adicionalmente, el panel se integra con la infraestructura de planta existente: tiene conexión directa a los sensores de la línea de producción y sincronización con el sistema MES para recibir y registrar datos automáticamente. También permite la exportación automática de reportes y el registro de incidencias en una base de datos central, facilitando las tareas de trazabilidad y cumplimiento regulatorio.

En cuanto a las tecnologías empleadas, el proyecto hace uso extensivo de Apache Spark para el procesamiento y análisis de datos a gran escala, lo cual fue fundamental dado el tamaño del histórico (orden de millones de filas) y la necesidad de preprocesamiento complejo. La selección de un modelo de Gradient Boosted Trees se basó en su buen desempeño en problemas de regresión y su capacidad para manejar variables heterogéneas (continuas y categóricas) sin mucho ajuste manual. El entrenamiento del modelo se realizó con Spark MLlib, aprovechando la capacidad distribuida para agilizar la optimización. Por otro lado, Streamlit proporcionó una vía rápida para construir el panel interactivo estilo SCADA, con la flexibilidad de una aplicación web moderna y la facilidad de despliegue que eso conlleva.



(accesible vía navegador web). Asimismo, la solución está diseñada pensando en su mantenibilidad: el pipeline de datos y el modelo pueden re-entrenarse periódicamente conforme se acumulen nuevos datos, y el panel está modularizado de forma que nuevos indicadores o gráficas se puedan incorporar con mínima intervención al código.

## **Resultados**

Los resultados obtenidos con el sistema desarrollado demuestran claras mejoras en el control de calidad del proceso cervecero. En términos cuantitativos de desempeño del modelo predictivo, se alcanzó un Error Cuadrático Medio (RMSE) menor a 0.15 en la predicción del puntaje de calidad, indicando que las estimaciones del modelo difieren muy poco del valor real de calidad en promedio. Adicionalmente, al categorizar los lotes entre aceptables y no aceptables, el sistema logró una precisión superior al 92%, lo que significa que más del 92% de las veces la IA acertó al predecir si un lote cumpliría con los estándares de calidad o no. Esta alta exactitud predictiva supera con creces la capacidad humana o métodos tradicionales, permitiendo confiar en las alertas generadas.

En el entorno de producción (simulado para pruebas), la implementación del sistema mostró beneficios operativos tangibles. El monitoreo continuo 24/7 mediante el panel de control permitió detectar el 99.7% de las anomalías críticas en el proceso de elaboración. En otras palabras, prácticamente todas las desviaciones significativas en parámetros (por ejemplo, un pH fuera de rango o una temperatura anómala) fueron identificadas por el sistema. Gracias a esto, el tiempo de respuesta ante condiciones fuera de lo normal mejoró dramáticamente: el personal ahora recibe alertas en menos de 3 minutos desde que ocurre una variación importante. Antes, dichas desviaciones podían pasar inadvertidas por periodos mucho más largos (hasta horas) si dependían únicamente de rondas de inspección manual. Con las alarmas tempranas, se pudo intervenir el proceso (ajustando temperatura, dosificando insumos, etc.) casi inmediatamente, evitando que el lote se arruine.

Otro resultado clave fue la reducción de los lotes rechazados y de los costos asociados al desperdicio. En las pruebas comparativas antes vs. después de la solución, se observó que la tasa de rechazo de lotes bajó

significativamente. Esto es consecuencia directa de la detección oportuna: al resolver las desviaciones durante la producción, los lotes terminan dentro de especificación y no necesitan ser descartados. Asimismo, al disminuir los rechazos, también se redujeron los costos de producción (menos materia prima desperdiciada, menos tiempo invertido rehaciendo lotes). Si bien las cifras exactas dependen del volumen de producción, la tendencia general fue clara: la cervecería pudo producir de forma más eficiente y con mayor calidad consistente. Incluso métricas de calidad global, como el puntaje promedio de calidad de los lotes, mostraron una mejora notable tras implementar el sistema (indicando menor variabilidad entre lotes y un control más estrecho del proceso).

En resumen, los resultados validan la hipótesis inicial: la solución basada en IA e IoT logró mejorar los indicadores de calidad y eficiencia del proceso cervecero. Se obtuvo un sistema confiable, rápido y preciso que eleva el estándar de control de calidad en planta, proporcionando a los ingenieros y operadores una herramienta poderosa para garantizar que cada cerveza producida cumpla con los más altos estándares.

## **Conclusiones**

En conclusión, el Sistema de Control de Calidad para Cervecería desarrollado demostró la eficacia de aplicar tecnologías de IoT, SCADA e Inteligencia Artificial en un entorno industrial cervecero. Se logró convertir un proceso de control de calidad mayormente reactivo e impreciso en uno proactivo y basado en datos, donde las decisiones pueden anticiparse a los problemas. Las ventajas del sistema son evidentes: la detección temprana de variaciones en parámetros críticos ha prácticamente eliminado las sorpresas desagradables en la calidad del producto, reduciendo desperdicios y mejorando la consistencia de la cerveza elaborada. Además, la automatización del monitoreo ha liberado a los operarios de estar revisando constantemente instrumentos, permitiéndoles enfocarse en acciones correctivas y mejoras de proceso en lugar de en la detección manual de fallos.

Cabe destacar que el éxito de este proyecto sienta las bases para futuras ampliaciones y mejoras. Algunas perspectivas a futuro y recomendaciones de mejora incluyen:

- **Calibración trimestral del modelo:** Re-entrenar el modelo de predicción cada cierto periodo (ej. cada 3 meses) incorporando los datos más recientes, para asegurar que la precisión predictiva se mantenga a medida que las recetas o condiciones de producción evolucionen. Esto equivale a una recalibración periódica del sistema de IA conforme cambie la realidad de la planta.
- **Migración a la nube:** Desplegar la infraestructura de datos y el panel de control en un entorno cloud. Esto permitiría una escalabilidad mucho mayor (poder almacenar y procesar años de datos sin limitaciones locales) y una alta disponibilidad. Un sistema en la nube facilitaría también integrar múltiples plantas cerveceras en la misma plataforma de control de calidad, habilitando comparaciones entre sitios y gestión centralizada.
- **Aplicación móvil:** Desarrollar un app móvil vinculada al panel de control, de modo que los supervisores puedan monitorear remotamente el proceso desde un teléfono o tableta. Esto brindaría mayor flexibilidad al equipo de calidad, que recibiría notificaciones push de alarmas críticas y podría revisar métricas clave en cualquier momento y lugar, sin tener que estar físicamente frente al panel en la fábrica.
- **Automatización avanzada del control:** Integrar más estrechamente el sistema de predicción con los controladores industriales (PLC/DCS) de la planta para cerrar completamente el lazo de control. En el futuro, ante una alerta de calidad, el sistema podría no solo notificar a humanos sino también ejecutar ajustes automáticos en tiempo real, por ejemplo, modificar la temperatura de un tanque o dosificar aditivos correctivos, para corregir la desviación sin demora. Esta automatización avanzaría el sistema hacia un control autónomo de calidad, minimizando aún más la posibilidad de error humano y optimizando la producción de manera dinámica.

En síntesis, el proyecto logró sus objetivos de demostrar cómo una cervecería puede beneficiarse de la convergencia de IoT, IA y sistemas de control industrial para elevar su control de calidad a estándares modernos. Las mejoras en precisión de predicción, tiempo de respuesta y reducción de rechazos representan no solo un beneficio económico, sino también un paso hacia la transformación digital de los procesos cerveceros. Con las mejoras propuestas a futuro, este sistema de control de calidad inteligente podría

escalarsse y volverse un componente esencial en cervecerías de mayor capacidad, impulsando una producción más eficiente, consistente y automatizada. Las herramientas desarrolladas (pipeline de datos, modelo GBT entrenado, panel Streamlit) sientan un precedente valioso para proyectos similares en la industria de bebidas y alimentos, donde la calidad del producto es crítica. En definitiva, la integración de tecnologías de vanguardia en la fábrica ha probado ser una estrategia ganadora para garantizar la excelencia en cada cerveza producida, posicionando a la cervecería un paso adelante en innovación y competitividad.