

DESAFÍO UNIVERSIDAD - EMPRESA

Esta necesidad tecnológica forma parte del Concurso de Proyectos de I+D+i y/o consultoría en colaboración Universidad – Empresa “Desafío Universidad Empresa” 2016 organizado por la Fundación Universidades y Enseñanzas Superiores de Castilla y León.

TÍTULO DE LA DEMANDA TECNOLÓGICA A RESOLVER

Referencia:

NT45

Título de la demanda tecnológica propuesta

Supercomputación para analítica predictiva de datos

Acrónimo:

SUPERCOMPUTACIÓN

Áreas de interés de la demanda tecnológica

(Principal) Tecnologías de la Información y Comunicación, Energía y Sostenibilidad

Resumen:

El volumen, la variedad y la velocidad de la información almacenada en las organizaciones están aumentando de manera significativa. El análisis inteligente de esos datos puede ser un factor diferenciador para las empresas que adopten esta técnica.

Para que las herramientas de analítica avanzada sean eficientes, estas han de soportar tecnologías de cálculo intensivo a través de la nube. Es pues necesario implementar paralelización en clusters de supercomputación.

DESCRIPCIÓN DE LA NECESIDAD DEMANDADA

1.- Descripción de la demanda tecnológica.

Para que una herramienta de analítica avanzada pueda competir con éxito en el mercado global, ningún estudio puede ser frenado por las limitaciones de un ordenador portátil o de sobremesa. Esto supone una necesidad tecnológica muy clara que la empresa necesita resolver cuanto antes: La paralelización en clústers de supercomputación para poder analizar grandes cantidades de datos a mucha más velocidad.

De esta manera, la interfaz del programa se alojará localmente en el ordenador del usuario, mientras que todo el procesamiento lógico se efectuará como un servicio web a través de la nube. La paralelización en clústers se realizará usando la biblioteca para programación en alto

rendimiento MPI. Esta hace uso de memoria distribuida, por lo que se eliminarán las restricciones de volumen y se multiplicará la velocidad.

En un primer momento, se podrán utilizar infraestructuras propias de la universidad o el centro de investigación. Una vez en producción, la empresa utilizará servidores propios de Microsoft Azure y/o Amazon AWS.

Con la conclusión de este proyecto, nuestra tecnología de minería de datos soportará tecnologías de big data y big computing, por lo que ni el tamaño del conjunto de datos ni el tiempo de procesamiento serán un problema.

Este sistema proporcionará un alto valor añadido a nuestros clientes, los analistas de datos, permitiéndoles obtener modelos predictivos inalcanzables con otros productos sustitutivos en el mercado.

2.- Antecedentes.

Nuestra empresa desarrolla tecnología disruptiva para la analítica predictiva usando métodos innovadores de aprendizaje profundo, más conocido como deep learning. El deep learning es un nuevo campo de la inteligencia artificial que combina las mejores técnicas de las redes neuronales y el análisis de datos para construir los modelos predictivos más potentes.

Algunas de las áreas en las que hemos aplicado estas técnicas con éxito son la ingeniería (optimización del rendimiento, mantenimiento predictivo,...), el marketing (segmentación de clientes, prevención del abandono,...) o la salud (diagnóstico médico, análisis de microarrays,...).

La empresa desarrolla tecnología propia de redes neuronales. Nuestro software ha participado en numerosos proyectos internacionales de I+D+i. Hoy en día, cuenta con miles de usuarios y es utilizada por importantes empresas.

Nuestros usuarios nos han transmitido su necesidad de analizar grandes cantidades de datos en un tiempo mínimo. Es pues imprescindible que nuestra tecnología implemente paralelización en clústers de supercomputación para cubrir las necesidades de nuestros clientes.

3.- Posibles enfoques del proyecto de investigación.

MPI, siglas en inglés de Message Passing Interface, es un protocolo de comunicación entre computadoras. Es el estándar para la comunicación entre los nodos que ejecutan un programa en un sistema de memoria distribuida. Este proyecto usará el estándar MPI para el procesamiento intensivo en supercomputadores.

Por otro lado, nuestra tecnología de redes neuronales ha sido escrita en C++. Por tanto, todos los desarrollos en el marco de este proyecto se realizarán con ese lenguaje de programación.

4.- Enfoques sin interés.

PALABRAS CLAVE: analítica predictiva, big data, supercomputación, paralelización

Si desea remitir una propuesta de solución tecnológica (proyecto de investigación y/o consultoría) deberá enviar el formulario de participación (ANEXO II), descargable en www.redtcue.es/, a una de las direcciones de correo electrónico que se indican en las bases del concurso antes del 15/06/2016.

[Acceso a información general del concurso](#)