

Deliverable D11 Item 8 – Daniel Carpio Camacho

A falta de solo un curso para acabar el grado de Ingeniería Informática-Ingeniería del Software, es el momento de empezar a pensar de manera mucho más específica a qué dedicarme, para que a partir de ahora pueda centrarme en las técnicas que se usaran en la vida laboral. Además, debo de empezar a pensar tanto en las empresas como los proyectos que en la actualidad se están llevando a cabo. En esto se basará este documento.

Hoy en día, la tecnología es algo inmensamente presente en nuestros días. E incluso más a nivel de negocios. Es por esto último que las empresas tienen miles de bytes de información de facturas de todos los clientes. Pero no solo tienen los diferentes productos que adquieren en compras y con qué frecuencia los obtienen, sino que de algunos tienen más datos, como datos personales o gustos y aficiones.

Todos estos datos, de por sí, no nos aportan gran cosa. Solo son ristas de números en enormes cantidades. Una persona no sacaría nada en claro simplemente ojeando esos números. Pero si metemos un ordenador de por medio, la cosa cambia. Es por ello por lo que existe una rama de la informática que se llama Minería de Datos que se encarga precisamente de eso. Estudiar y analizar esos datos para poder hallar patrones o conjuntos que permitan sacar información útil que las empresas puedan usar.

Es precisamente esto a lo que me gustaría dedicarme laboralmente. Aprender diferentes técnicas de aprendizaje automático que nos permitan sacar información de todos esos datos.

¿Qué estoy haciendo actualmente? En la actualidad, gracias a numerosas ofertas que nos ofrece la facultad, pude presentarme como alumno interno para el estudio de estas técnicas, ya que había una oferta. Tras un cuatrimestre, diferentes profesores me han enseñado la teoría que se esconde detrás de las técnicas y a utilizar algunos programas que ayudan a utilizar dichas técnicas. Programas como Weka o KNIME son algunos ejemplos de estos programas, aunque existen otros programas también muy usados y bastante interesantes de utilizar.

Debido a que ya tenía que formarme en el lenguaje de programación Python para la asignatura de Inteligencia Artificial que, dicho sea de paso, también toca algunos de los aspectos vistos en la minería de datos, aprovecharé para aprender por mi cuenta a utilizar una conocida librería llamada scikit-learn que implementa estas técnicas. Igualmente, aprenderé acerca del lenguaje de programación R que, debido a su calidad a la hora de trabajar con muchos datos, lo hace también idóneo para trabajar en el mundo de la minería de datos. Debido a que los programas como Weka o la librería scikit-learn son de software libre, gran cantidad de gente los usa y, a su vez, existe una gran cantidad de documentación de las mismas por Internet.

Saber utilizar distintos programas ayuda a incorporarse a la manera de trabajar que tenga el futuro trabajo que realice, ya que distintas técnicas pueden variar bastante de la una a la otra. Si nos fijamos en lo descrito anteriormente, se ve como uno usa programas mientras que otro se hace en código mediante el uso de librerías. Esto supone que el cambio de uno a otro suponga un cambio grande, por lo que previene estar preparado para acomodarse rápidamente en uno de los dos.

Además, hay numerosas investigaciones para conseguir encontrar mejores técnicas que las actuales. Principalmente se basan en encontrar los que den un mejor grado de acierto, ya que con las técnicas actuales puede que cueste encontrar información con pocos errores. A su vez, se intenta encontrar técnicas que permitan obtener estos resultados en un tiempo más rápido. En función de la cantidad de datos que se use, puede ser que el tiempo de ejecución sea muy elevado, pudiendo ser bastante molesto si además de tardar mucho, el resultado no sea el esperado.

Un punto importante para conseguir alcanzar mis objetivos es no dejar los estudios en el grado. Principalmente, tengo planeado la realización de uno de los másteres que la propia facultad nos ofrece. En un futuro próximo, tengo planeado quedar con algún profesor de los que me enseñó sobre la minería de datos que me pueda aconsejar más en profundidad sobre dichos másteres y sus asignaturas, así que de momento no puedo entrar mucho en detalle. Sí que es cierto que ya me comentó que la facultad ofrece buenas opciones para el estudio de este tema, aunque bien es cierto que en el grado no nos ofrece mucha información relativa a estas técnicas, sobre todo si tenemos en cuenta el gran impacto que tiene hoy en día.

Paralelamente a todo esto, está la importancia de los idiomas, principalmente del inglés. En la actualidad, tengo el título B1 por el Trinity College London, aunque me estoy preparando para el título B2 de Cambridge. La preparación en el idioma inglés lo considero algo obligatorio, es por ello por lo que más adelante me prepare para niveles superiores.

Es importante tener en cuenta que numerosas empresas buscan a gente que estén estudiando estos temas para ofrecer becas. Algunas de estas empresas son como Hitachi o Endesa, las cuales no hace más de un mes que ofrecieron becas de ayudas al estudio del Big Data.

Estas empresas conocen las ventajas de todo lo comentado al principio, y buscan a gente con estudios para incorporarse a trabajar dentro de sus plantillas para intentar aprovechar toda la información que puedan obtener de sus datos que, al ser empresas grandes, se puede obtener bastante.