

Towards an Integrated Educational Platform to Help Assess Capstone Project Teams

Andres Neyem ¹[0000-0002-5734-722X], Juan Diaz-Mosquera ²[0000-0002-6246-2126] and Jose Ignacio Benedetto ³[0000-0003-2918-3640]

Pontificia Universidad Católica de Chile, Santiago, Chile
¹aneyem@uc.cl, ²jddiaz4@uc.cl, ³jibenedetto@uc.cl

Abstract. Capstone courses offer a favorable environment for the development of student's skills through an approach with theoretical and practical components. However, proper coordination between the students, stakeholders and academic team is very difficult to achieve. The lack of proper tools for addressing this issue along with time constraints often hinder the students from attaining the desired course outcomes. This paper presents an integrated environment to conduct capstone courses in Software Engineering. We make a review of the framework for best practices applied in these courses and describe SharedBoard, an educational tool used as a central supporting for managing each student project and as a source for our data analysis. We propose a series of Learning Analytics to verify the alignment of student activities with learning outcomes in management skills. Analytics are shown to validate the abilities of the students for planning, scheduling and tracking their projects. Finally, conclusions and future works are discussed in this paper.

Keywords: Learning Analytics, Capstone, Software Engineering.

1 Introducción

A través de los cursos basados en proyectos, estudiantes en áreas de ciencias de la computación tienen la oportunidad de ejercitar las habilidades profesionales que necesitarán después de la universidad. Proyectos con restricciones de la vida real basados en las necesidades y especificaciones de clientes actuales son particularmente relevantes ya que son lo más efectivo para reducir la brecha entre la industria y la academia.

Para las universidades, enseñar cursos basados en proyectos es un gran desafío e involucra fuertes responsabilidades para proveer ambientes donde los estudiantes desarrollen sus habilidades profesionales [3]. El curso en la Universidad Católica de Chile (UC), es un curso capstone diseñado alrededor de esta visión. La intención de este curso es proveer una experiencia capstone que integre el conocimiento ganado de todos los cursos previos en el currículo trabajando en un proyecto con desafíos realistas.

A medida que los proyectos avanzan, la complejidad que rodea a estos incrementa sustancialmente. Gráficas, estadísticas, métricas y reportes en general se vuelven muy importantes para entender que está pasando en los proyectos y ayudar en su administración. En el caso de los cursos capstone, esto se vuelve incluso más importante ya que

permite que el equipo académico siga el proceso de aprendizaje de los estudiantes. Algunas investigaciones se centran en métricas y estrategias para determinar el nivel de aplicación de habilidades de administración de proyectos y de metodologías de desarrollo de software como Scrum, XP, Kanban entre otras [2, 11]. A pesar de que estas propuestas presentan un enfoque fresco en la misma dirección de nuestro trabajo, estos carecen de un ecosistema unificado para soportar experiencias de enseñanza y aprendizaje.

Este trabajo propone un análisis de la información extraída del curso capstone en la UC. Esto es hecho usando los datos que fueron almacenados en la herramienta educativa SharedBoard. Con analíticas de aprendizaje, se busca validar la correcta aplicación de habilidades a nivel de administración y desarrollo de software. Estos resultados permiten a profesores y estudiantes mejorar las experiencias capstone ya que permiten identificar los aspectos más relevantes de las actividades realizadas por estos últimos, además proveen una vista del comportamiento de ellos en los proyectos. Mas allá, estos resultados ayudan a los equipos académicos a tomar acción en caso de situaciones inesperadas.

La siguiente sección provee una breve descripción del marco de trabajo del curso capstone junto con sus muchas ventajas. La sección 3 presenta a la herramienta SharedBoard como un sistema móvil basado en la nube, junto con sus variadas características y como estas soportan al marco de trabajo previamente mencionado. Esta sección también ilustra el potencial del enfoque previo usando analíticas generadas por SharedBoard y aplicando técnicas de Ciencias de Datos para la Ingeniería de Software [9] con el fin de minar y extender la información de esta plataforma. Finalmente, la sección 4 presenta conclusiones y trabajo futuro.

2 Marco de trabajo capstone: Mejores Practicas

Proyectos con situaciones de la vida real se han vuelto comunes en los cursos capstone en Ingeniería de Software en varias universidades debido a su valor pedagógico en cerrar la brecha entre los mundos de la industria y la academia [4]. Aun así, estos tipos de proyectos poseen varios desafíos para los equipos académicos [12].

Después de muchos años de enseñar Ingeniería de Software Empírica, observaciones han permitido a Neyem et al. desarrollar un marco de trabajo de mejores prácticas para conducir cursos capstone en Ciencias de la Computación que correctamente aborden todos los desafíos nombrados anteriormente [12]. Esto involucra un enfoque basado en proyectos para mejorar la experiencia de aprendizaje también como la integración de varias herramientas que de gran manera facilitan la administración de proyectos, el seguimiento, la trazabilidad y la comunicación entre las diferentes partes.

La meta de este marco de trabajo es familiarizar a los estudiantes con las mejores prácticas para desarrollar software de alta calidad a la par con los estándares comerciales en un ambiente no diferente al encontrado en las industrias de software más prestigiosas de la actualidad. Este marco de trabajo general es alineado con una serie de logros de aprendizaje que buscan que los estudiantes sean hábiles para trabajar en grupos

con el objetivo de especificar requerimientos, diseñar, probar e implementar software solicitado por clientes reales, desarrollar habilidades de administración de proyectos y mejorar las habilidades de comunicación.

La idea general del marco de trabajo de mejores prácticas consiste en 5 fases: trabajo previo, introducción y tres ciclos separados de desarrollo de software. La fase de trabajo previo deberá enfocarse en la selección de proyectos; la fase de introducción deberá hacer énfasis en la enseñanza de la metodología del curso y los principios de ingeniería de software que pueden ser aplicados; finalmente los ciclos o “sprints” de desarrollo de software abundan en las prácticas actuales de desarrollo (levantamiento de requerimientos, diseño, programación, pruebas y puesta en marcha del software). Una descripción más profunda de cada una de estas fases junto con objetivos, sugerencias, ejemplos de implementación, desafíos comunes y dificultades, está disponible en [12].

3 Herramienta de Software Educativo y Analíticas de Aprendizaje

Para que un curso capstone progrese fluidamente, una precisa administración y seguimiento de proyectos es requerida tanto para los grupos de estudiantes (para coordinar efectivamente sus esfuerzos) como para el equipo académico (para propósitos de evaluación). Esto se traduce a un gran desafío para los líderes de los grupos de proyectos quienes usualmente requieren completa dedicación a la administración de sus proyectos asignados, y en una inmanejable tarea para el equipo académico que necesita mantener el seguimiento de cada proyecto con el fin de proveer una evaluación justa y una valiosa retroalimentación.

SharedBoard fue creado con el objetivo de proveer una solución integral para el seguimiento de proyectos compatible con la escuela. SharedBoard es un Sistema Móvil basado en la Nube (CMS, sigla en inglés) que funciona como un espacio de trabajo virtual compartido. Los espacios de trabajo compartido buscan soportar tareas cooperativas ofreciendo un espacio virtual donde la información puede ser compartida e intercambiada. Por lo general, estas herramientas se centran en el intercambio de información en el sentido de autoría cooperativa, comentando y anotando documentos compartidos como una actividad grupal.

SharedBoard consiste en un componente web y móvil. El componente móvil se comunica con el componente web a través de una API RESTful para sincronizar datos; sin embargo, tiene su propia base de datos local para habilitar la operación fuera de línea. Su objetivo principal es recopilar información personalizada sobre la rutina de trabajo y las contribuciones de los miembros del grupo, y notificarles sobre cualquier cambio de estado relevante de las tareas que les han asignado.

SharedBoard busca cubrir todas las necesidades básicas de los proyectos modernos de software ágil. Estas actividades incluyen soporte para metodologías formales, gestión de requisitos, gestión de equipos, pruebas de software, colaboración y documentación [16]. Las funcionalidades principales se centran en "representaciones de tareas" basadas en la filosofía del tablero Kanban. Los tableros Kanban son muy populares en

la industria del software hoy en día. Se ha demostrado empíricamente que proporcionan una mayor motivación y control de la actividad del proyecto debido a su simplicidad inherente [6]. Si bien las tareas registradas en un Kanban no son más que trabajos granulares, SharedBoard lleva más lejos esta idea al permitir a los participantes enriquecerlas al vincularlas a otros módulos del proyecto. Como tal, el tablero Kanban sirve como un centro de comunicación para toda la aplicación.

Por otro lado, las Analíticas de Aprendizaje son instrumentos para medir, reunir, analizar y mostrar información acerca del comportamiento de los estudiantes en un proceso de aprendizaje [7]. En cuanto a su aplicación a los cursos capstone, hay algunos estudios [10, 13, 14] que presentan diferentes enfoques para interpretar y visualizar información de los estudiantes. Sin embargo, estos estudios carecen de una herramienta educativa integrada para respaldar el proceso de aprendizaje en un solo punto, como lo hace SharedBoard. Una vez dicho esto, se presentan una serie de analíticas y se proponen técnicas de minería de datos y procesos para analizar y ampliar la información recopilada de los cursos capstone.

3.1 Administración Visual del Trabajo de Equipo

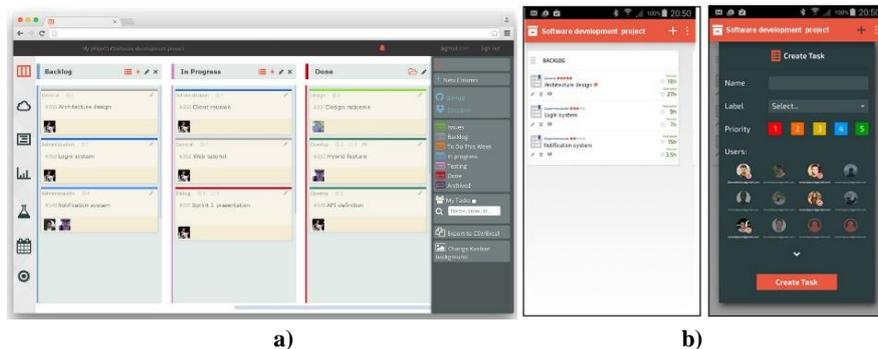


Fig. 1. SharedBoard: a) Vista Web del Kanban. b) Vista Móvil del Kanban.

Las Figuras 1a y 1b muestran el tablero Kanban como un espacio virtual compartido que ofrece a todo el equipo una visión global del proyecto, junto con todas las tareas en las que están trabajando actualmente. Los miembros del equipo pueden asignar tareas a uno o más compañeros, actualizar su estado, compartir sus opiniones a través de comentarios y categorizarlos a través de etiquetas. El tablero Kanban es la forma más efectiva de visualizar el flujo de trabajo y permite que el equipo de estudiantes tenga control sobre lo que se debe hacer y sobre lo que se está desarrollando actualmente.

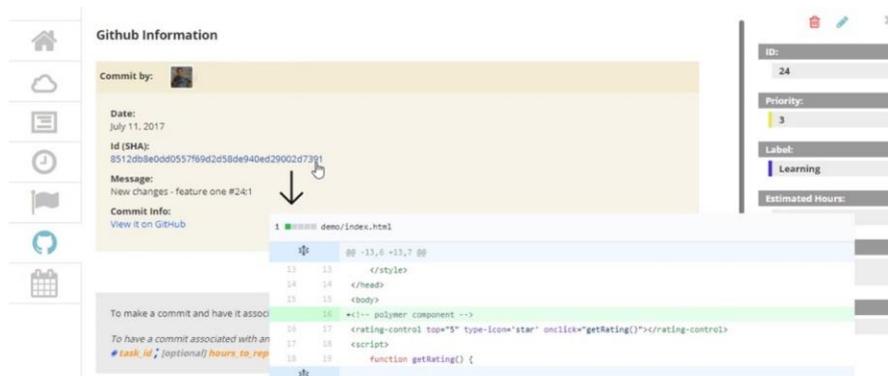


Fig. 2. SharedBoard integrado con GitHub.

Específicamente para el desarrollo de software, garantizar la trazabilidad entre una tarea y su código asociado se convierte en un factor clave para mantener un seguimiento más amplio de las actividades del proyecto. La Figura 2 muestra cómo se integra SharedBoard con GitHub para cumplir este objetivo. La integración consiste inicialmente en enlazar por medio de un asistente un tablero Kanban de SharedBoard con un proyecto en GitHub, posteriormente, al agregar un comentario en cada commit con un formato preestablecido (id tarea + comentario) se creará un vínculo que permitirá hacer la trazabilidad de los commits en SharedBoard.

3.2 Administración Visual de Requerimientos

Requirements

Here you can list all the requirements of your project. You can select from default templates or create one just for your project.

You can link your tasks with the requirements.

Note: you can only edit the requirements of your tasks.

You can select to see requirements by groups or hierarchies.

Search Requirement...

Requirement Groups Requirements Hierarchy

User Story Propio
Control de identidad
Acceder a la aplicación
Manejo de abuso de cuentas
Descarga selectiva de Atlas
Asignación cuerpo docente de una sección
Anotaciones

New Requirement Task-Requirement Matrix Requirement Templates Generate PDF

Fig. 3. Módulo de requerimientos: Creación.

En SharedBoard, los estudiantes tienen la capacidad de realizar un seguimiento de los requisitos del proyecto. Estos pueden asociarse con tareas registradas en el tablero Kanban. Es posible crear historias de usuarios o casos de uso para reflejar los requisitos deseados por las partes interesadas (ver Figura 3), y también cada requisito puede tener una lista de verificación para seguir su progreso (ver Figura 4).

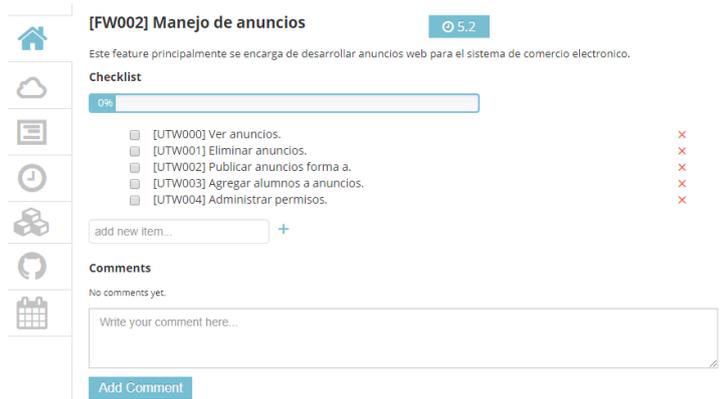


Fig. 4. Lista de verificación de requerimientos.

3.3 Reportes de Trabajo en Equipo

SharedBoard proporciona a sus usuarios diferentes tipos de informes. En la Figura 5 hay una visualización de datos para mostrar un resumen de las horas de trabajo reportadas por los estudiantes en el desarrollo de sus actividades. Además, se pueden ver estas horas reportadas filtradas por categorías como por ejemplo funcionalidades, pruebas, análisis, etc.

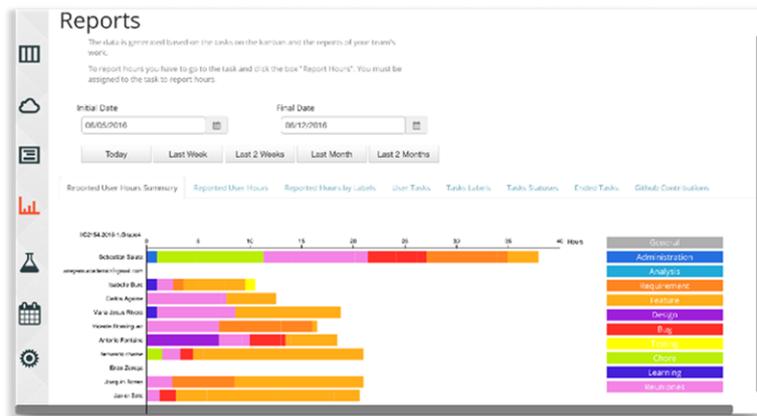


Fig. 5. Resumen de las horas reportadas por los usuarios en SharedBoard.

SharedBoard también permite visualizar las tareas catalogadas como finalizadas. Esta opción permite a los usuarios filtrar estas tareas por atributos tales como horas o cantidad de tareas. Además, es posible ir en detalle en cada categoría haciendo clic en las barras.

3.4 Trabajo en Equipos: Análisis Basado en Minería de Procesos

La Minería de Procesos (MP) es un área de investigación con un amplio alcance que proporciona a las esferas académica y empresarial los métodos y herramientas necesarios para encontrar aspectos significativos y relevantes de los procesos organizacionales [17]. MP es una dimensión de la Ciencia de los Datos que emplea esta trabajo a través de un enfoque de Ingeniería de Software Empírica para amplificar el conocimiento obtenido de SharedBoard de una manera orientada a procesos.

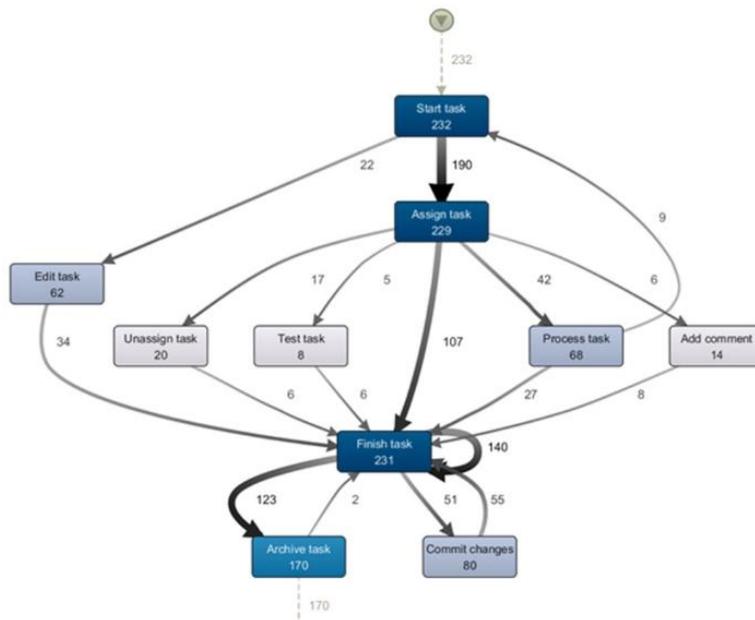


Fig. 6. Modelo MP obtenido de un proyecto del curso capstone.

Hoy en día, hay varios enfoques, métodos, algoritmos y herramientas para MP; DISCO [8] y ProM [5] son las dos herramientas más utilizadas, con fines comerciales y académicos, respectivamente. Estas herramientas incluyen muchos algoritmos y utilidades para ayudar en las tres fases clave en esta área de investigación de MP: descubrimiento, verificación de conformidad y mejoramiento [17]. Se han implementado las primeras dos fases. A través del descubrimiento, es revelado el flujo desarrollado de un proceso. Luego es usada la conformidad para determinar el grado de afinidad que el flujo descubierto tiene en contraste con un modelo propuesto por SharedBoard (este es, un flujo de trabajo con actividades tales como asignar recursos, desarrollar tareas, probar tareas, etc.).

Este análisis ha sido llevado a cabo con estudiantes del curso capstone en Ingeniería de Software de la UC durante los dos semestres académicos del 2017. En cada semestre han participado alrededor de 50 estudiantes formándose 5 equipos en cada periodo. Los grupos han trabajado en proyectos que van desde el desarrollo de plataformas tipo B2C (Business to consumers) hasta dispositivos embebidos que involucran programación en Arduino. Normalmente los estudiantes siguen el proceso de identificar tareas, desarrollarlas y cerrarlas teniendo en cuenta que cada proceso se puede extender o variar.

Con MP, se obtuvieron modelos de procesos (ver Figura 6) que ayudaron a comprender el flujo de trabajo dentro de cada proyecto del curso. De acuerdo con el grosor de la línea, se puede establecer el nivel de interacción entre las actividades; el color de la caja representa la frecuencia de cada actividad. Esta vista permite a los estudiantes y al equipo académico identificar situaciones fuera de lo común (flujos omitidos, cuellos de botella, etc.) para tomar las medidas adecuadas; y también para identificar puntos fuertes en comportamientos de gestión (correcta asignación de tareas, flujos de trabajo completos, etc.). Del mismo modo, MP llevo a identificar las interacciones entre los miembros del equipo, en la Figura 7 se puede ver un gráfico que lo representa.

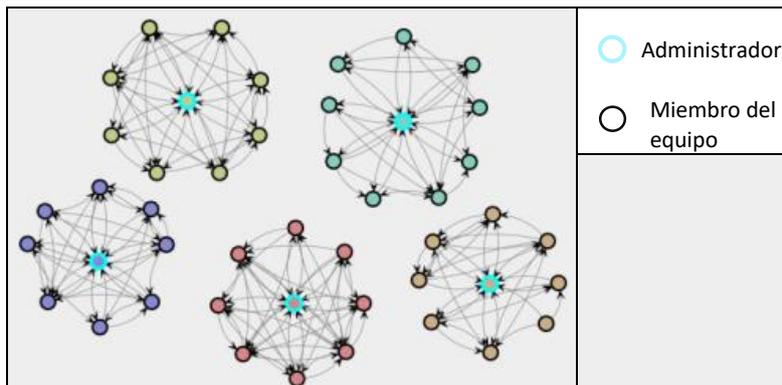


Fig. 7. Identificando la interacción de los equipos usando MP.

Los nodos en ese gráfico muestran un adecuado flujo entre una figura de administrador (nodo cian) y los miembros de su equipo, en otras palabras, se puede corroborar que los administradores funcionan como un eje central en cada uno de los proyectos.

3.5 Análisis Basado en Minería de Datos de las Actividades de los Estudiantes

La Minería de Datos (MD) [1] es el proceso de descubrir conocimiento de diferentes tipos de repositorios utilizando técnicas específicas de análisis de información. En SharedBoard, es utilizada la técnica TF-IDF [15] que consiste en un método para ponderar la relevancia de las palabras en un grupo de documentos o registros. Esta técnica tiene en cuenta las frecuencias de las palabras y el número de veces que estas fueron utilizadas en diferentes documentos.

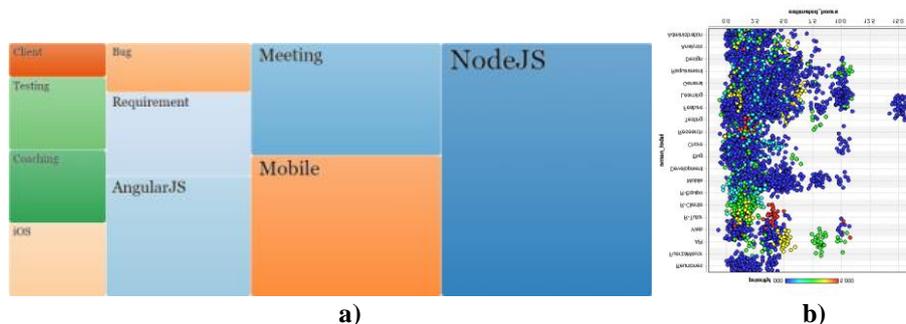


Fig. 8. a) Treemap generado al minar un proyecto del curso capstone, b) Grafica de dispersión con el nivel de prioridad de las actividades

Este trabajo presenta un análisis aplicando técnicas de MD que busca proporcionar a los usuarios con diferentes formas de visualizar lo que sucede dentro de los proyectos. La Figura 8a muestra un treemap que se generó después de minar con TF-IDF varios registros de tablas de SharedBoard con información significativa (tareas, subtareas y commits). Por otro lado, la Figura 8b presenta un gráfico de dispersión para descubrir cómo los estudiantes priorizan sus actividades de acuerdo con las categorías y horas estimadas. La idea es brindar a los usuarios una nueva visión de los aspectos relevantes de sus proyectos.

4 Discusiones y Conclusiones

El enfoque antes mencionado ha mostrado ser una alternativa adecuada para conducir cursos capstone en Ingeniería de Software. La información extraída de estos cursos en la UC durante 1 año ha llevado a este trabajo a obtener resultados significativos. Los estudiantes reportan estar teniendo experiencias enriquecedoras desarrollando su proyecto. Las habilidades de gestión apropiadas se reflejan en la estrategia de división de tareas por categoría y en el registro de requisitos para el seguimiento de estas. Por otro lado, las técnicas de MP han ayudado a evaluar las habilidades del trabajo en equipo de los estudiantes a través de análisis visuales; MD ha apoyado la identificación de los temas más discutidos para cada fase en los cursos capstone.

Como trabajo futuro, este trabajo quiere extender el análisis de la información de los estudiantes usando más técnicas de Aprendizaje de Maquina. Se espera evaluar e implementar Sistemas de recomendación para mejorar la experiencia de los usuarios. Además, se planea extraer y analizar los repositorios de Github de todos los proyectos en el curso capstone para obtener información más valiosa que podría enriquecer enfoques futuros.

5 Agradecimientos

Agradecimientos al Departamento de Ciencias de la Computación, a la Unidad de Educación en Ingeniería de la Escuela de Ingeniería y a VRA FondeDOC por el apoyo

continuo de este curso. Finalmente, se agradece a todos los estudiantes del curso capstone IIC2154 que participaron en este proyecto de investigación educativa.

Referencias

- [1] Aggarwal C. C.: *Data mining: the textbook*. Springer (2015).
- [2] Alperowitz, L.; Dora D.; Bernd B.: Metrics in Agile project courses. *Proceedings of the 38th International Conference on Software Engineering Companion*. ACM (2016).
- [3] Chamillard, A. T.; Kim A. B.: The software engineering capstone: structure and tradeoffs. *ACM SIGCSE* (2002).
- [4] Chase, J. D.; Uppuluri, P.; Lewis, T.; Barland, I.; Pittges, J.: Integrating Live Projects Into Computing Curriculum. *Proceedings of the 46th ACM Technical Symposium on Computer Science Education*. ACM (2015).
- [5] Claes, J.; Geert P.: Process mining and the ProM framework: an exploratory survey. *International Conference on Business Process Management*. Springer Berlin Heidelberg (2012).
- [6] Ikonen, M.; Pirinen, E.; Fagerholm, F.; Kettunen, P.; Abrahamsson, P.: "On the impact of kanban on software project work: An empirical case study investigation." *Engineering of Complex Computer Systems (ICECCS), 2011 16th IEEE International Conference on*. IEEE (2011).
- [7] Ferguson, R.: Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, Vol. 4, No 5-6, pp. 304-317 (2012).
- [8] Günther, C. W., and Rozinat A.: Disco: Discover Your Processes. *BPM (Demos)*, Vol. 940, pp. 40-44 (2012).
- [9] Menzies, T.; Williams, L.; Zimmermann, T.: *Perspectives on Data Science for Software Engineering*. Morgan Kaufmann (2016).
- [10] Mittal, M.; Sureka, A.: Process mining software repositories from student projects in an undergraduate software engineering course. *Companion Proceedings of the 36th International Conference on Software Engineering*. ACM (2014).
- [11] Mohan, S.; Chenoweth S.; S. Bohner. S.: Towards a better capstone experience. *Proceedings of the 43rd ACM technical symposium on Computer Science Education*. ACM (2012).
- [12] Neyem, A.; Benedetto, J. I.; Chacon, A. F.: Improving software engineering education through an empirical approach: lessons learned from capstone teaching experiences. *Proceedings of the 45th ACM technical symposium on Computer science education*. ACM (2014).
- [13] Petkovic, D.: Using Learning Analytics to Assess Capstone Project Teams. *IEEE Computer*, Vol. 49, No 1, pp. 80-83 (2016).
- [14] Poncin, W.; Serebrenik, A.; van den Brand, M.: Mining student capstone projects with FRASR and ProM. ACM, pp. 87-96 (2011).
- [15] Ramos, J.: Using tf-idf to determine word relevance in document queries. *Proceedings of the first instructional conference on machine learning*. (2003).
- [16] Uickey, N.; Suman, U.: An empirical study to design an effective agile project management framework. *Proceedings of the CUBE International Information Technology Conference*. ACM, pp. 385-390 (2012).
- [17] Van der Aalst, W.: *Process Mining: Data Science in Action*. Springer (2016).