



UNIVERSIDAD ANDINA DEL CUSCO

FACULTAD DE INGENIERÍA Y ARQUITECTURA

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS



TESIS

SISTEMA DE DESCRIPCIÓN DE DESERCIÓN UNIVERSITARIA EN LA UNIVERSIDAD ANDINA DEL CUSCO HACIENDO USO DEL DATA MINING Y LA METODOLOGÍA CRISP – DM

LÍNEA DE INVESTIGACIÓN:

TECNOLOGÍAS DE LA INFORMACIÓN

Presentado por:

Bach. Reime Figueroa Neyl

Bach. Valdeiglesias Romero Americo

Para optar al título de:

Ingeniero de Sistemas

Asesor: Mg. Vargas Vera Lizet

Co – Asesor: Mg. Pilar Hidalgo León

CUSCO - PERÚ

2019



Resumen

La Universidad Andina del Cusco al igual que la mayor parte de universidades en Latino América presenta índices de deserción universitaria por parte de sus alumnos lo cual representa un impacto negativo tanto para la Universidad Andina del Cusco como para los alumnos que desertaron de su escuela profesional, la Universidad Andina del Cusco actualmente no cuenta con un plan de contingencia frente a este problema y no tiene un concepto definido sobre cuando un alumno es desertor o no. En el presente trabajo de investigación de tesis el sistema de descripción se hará bajo la aplicación de la minería de datos y con la metodología CRISP – DM, que con la ayuda de la información proporcionada por la Universidad Andina del Cusco se realizó un análisis de información de los alumnos provenientes de las bases de datos de la Universidad Andina del Cusco (Servicios Académico), con dicha información se ha podido conocer las características sociales, económicas y académicas que describen a un alumno desertor, tomando como tiempo de estudio del semestre 2010 – I (Ingreso de los alumnos) hasta el semestre 2015 – II dando un periodo de gracia de cuatro semestres para que el alumno pueda egresar de su escuela profesional. Este periodo de gracia se da bajo la investigación de Díaz (2008) “Considera la inactividad de un alumno en su casa de estudio por tres semestres académicos”, en este caso se ha dado cuatro semestres debido a la poca cantidad de alumnos egresados de la Universidad Andina del Cusco, es por esto que para obtener un análisis adecuado se determinó ese tiempo de estudio. Los alumnos que no han concluido sus estudios universitarios, serán considerados como desertores, los alumnos que usaremos de muestra son pertenecientes a las escuelas profesionales de Ingeniería de Sistemas, Ingeniería Civil, Contabilidad y Turismo. Este sistema de descripción permitirá conocer las características de los alumnos desertores y con este conocimiento la Universidad Andina del Cusco en un futuro implemente estrategias necesarias que eviten que dichos alumnos deserten de su escuela profesional.

PALABRAS CLAVES: Deserción estudiantil, metodología CRISP-DM, clustering, asociación, clasificación, KDD, base de datos, tecnología de información, minería de



datos, proceso de conocimiento, descubrimiento de conocimientos en bases de datos, EM, k-medianas, WEKA, Python, pandas, estadística.

Abstract

The Andean University of Cusco, like most universities in Latin America, presents dropout rates on the part of its students, which represents a negative impact both for the Andean University of Cusco and for students who dropped out of their professional school, The Universidad Andina del Cusco does not currently have a contingency plan for this problem and does not have a definite concept of when a student is a deserter or not. In the present work of thesis research the description system will be done under the application of data mining and with the CRISP - DM methodology, which with the help of the information provided by the Universidad Andina del Cusco, an analysis of information was carried out of students from the databases of the Universidad Andina del Cusco (Academic Services), with this information has been able to know the social, economic and academic characteristics that describe a student deserter, taking as the time of study of the semester 2010 - I (Income of students) up to the semester 2015 - II giving a period of grace of four semesters so that the student can graduate from his professional school. This period of grace is given under the research of Díaz (2008) "It considers the inactivity of a student in his house of study for three academic semesters", in this case there have been four semesters due to the small number of students graduated from the Universidad Andina del Cusco, this is why to obtain an adequate analysis that study time was determined. Students who have not completed their university studies will be considered as deserters, the students that we will use as sample are from the professional schools of Systems Engineering, Civil Engineering, Accounting and Tourism. The present model of data mining will allow the Andean University of Cusco to know these characteristics of the deserting students and with this knowledge this university can in the future implement necessary strategies that prevent these students from deserting their professional school.