Proyecto deserción universitaria

¿Es posible predecir en una etapa temprana si un alumno tiene riesgo de desertar o no?

Esta pregunta puede ser muy común en sistemas educacionales donde trabajan con una cantidad inmensa de alumnado, en los cuales esta problemática cada vez se hace mas recurrente debido a la gran cantidad de jóvenes que por múltiples acontecimientos personales o sociales como la pandemia no se han adaptado al entorno universitario.

Para este proyecto utilizaré una base de datos del Instituto politécnico de Porto Alegre donde se ha recopilado un conjunto de datos a partir de una institución de educación superior (obtenido de varias BDD inconexas) relacionado con estudiantes matriculados en diferentes títulos universitarios como agronomía, diseño, educación, enfermería, periodismo, administración, servicio social y tecnologías. El conjunto de datos incluye información conocida al momento de la inscripción de los estudiantes (ruta académica, demografía y factores socioeconómicos) y el rendimiento académico de los estudiantes al final del primer y segundo semestre.

El conjunto de datos se creó en un proyecto que tiene como objetivo contribuir a la reducción de la deserción y el fracaso académico en la educación superior, mediante el uso de técnicas de aprendizaje automático para identificar a los estudiantes en riesgo en una etapa temprana de su trayectoria académica, de modo que se puedan diseñar estrategias para concentrar el apoyo académico en estos estudiantes.

El problema se formula como una tarea de clasificación de tres categorías:

**Abandono**

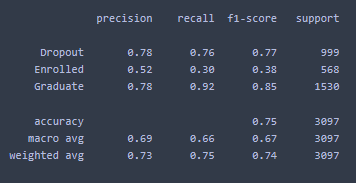
**Matriculado**

**Graduado**

Descripción base de datos:

|  |  |
| --- | --- |
| ATTRIBUTE NAME | DESCRIPTION |
| Marital status | 1 – single 2 – married 3 – widower 4 – divorced 5 – facto union 6 – legally separated |
| Application mode | 1 - 1st phase - general contingent 2 - Ordinance No. 612/93 5 - 1st phase - special contingent (Azores Island) 7 - Holders of other higher courses 10 - Ordinance No. 854-B/99 15 - International student (bachelor) 16 - 1st phase - special contingent (Madeira Island) 17 - 2nd phase - general contingent 18 - 3rd phase - general contingent 26 - Ordinance No. 533-A/99, item b2) (Different Plan) 27 - Ordinance No. 533-A/99, item b3 (Other Institution) 39 - Over 23 years old 42 - Transfer 43 - Change of course 44 - Technological specialization diploma holders 51 - Change of institution/course 53 - Short cycle diploma holders 57 - Change of institution/course (International) |
| Application order | Application order (between 0 - first choice; and 9 last choice) |
| Course | 33 - Biofuel Production Technologies 171 - Animation and Multimedia Design 8014 - Social Service (evening attendance) 9003 - Agronomy 9070 - Communication Design 9085 - Veterinary Nursing 9119 - Informatics Engineering 9130 - Equinculture 9147 - Management 9238 - Social Service 9254 - Tourism 9500 - Nursing 9556 - Oral Hygiene 9670 - Advertising and Marketing Management 9773 - Journalism and Communication 9853 - Basic Education 9991 - Management (evening attendance) |
| Daytime/evening attendance | 1 – daytime 0 - evening |
| Previous qualification | 1 - Secondary education 2 - Higher education - bachelor's degree 3 - Higher education - degree 4 - Higher education - master's 5 - Higher education - doctorate 6 - Frequency of higher education 9 - 12th year of schooling - not completed 10 - 11th year of schooling - not completed 12 - Other - 11th year of schooling 14 - 10th year of schooling 15 - 10th year of schooling - not completed 19 - Basic education 3rd cycle (9th/10th/11th year) or equiv. 38 - Basic education 2nd cycle (6th/7th/8th year) or equiv. 39 - Technological specialization course 40 - Higher education - degree (1st cycle) 42 - Professional higher technical course 43 - Higher education - master (2nd cycle) |
| Previous qualification (grade) | Grade of previous qualification (between 0 and 200) |
| Nacionality | 1 - Portuguese; 2 - German; 6 - Spanish; 11 - Italian; 13 - Dutch; 14 - English; 17 - Lithuanian; 21 - Angolan; 22 - Cape Verdean; 24 - Guinean; 25 - Mozambican; 26 - Santomean; 32 - Turkish; 41 - Brazilian; 62 - Romanian; 100 - Moldova (Republic of); 101 - Mexican; 103 - Ukrainian; 105 - Russian; 108 - Cuban; 109 - Colombian |
| Mother's qualification | 1 - Secondary Education - 12th Year of Schooling or Eq. 2 - Higher Education - Bachelor's Degree 3 - Higher Education - Degree 4 - Higher Education - Master's 5 - Higher Education - Doctorate 6 - Frequency of Higher Education 9 - 12th Year of Schooling - Not Completed 10 - 11th Year of Schooling - Not Completed 11 - 7th Year (Old) 12 - Other - 11th Year of Schooling 14 - 10th Year of Schooling 18 - General commerce course 19 - Basic Education 3rd Cycle (9th/10th/11th Year) or Equiv. 22 - Technical-professional course 26 - 7th year of schooling 27 - 2nd cycle of the general high school course 29 - 9th Year of Schooling - Not Completed 30 - 8th year of schooling 34 - Unknown 35 - Can't read or write 36 - Can read without having a 4th year of schooling 37 - Basic education 1st cycle (4th/5th year) or equiv. 38 - Basic Education 2nd Cycle (6th/7th/8th Year) or Equiv. 39 - Technological specialization course 40 - Higher education - degree (1st cycle) 41 - Specialized higher studies course 42 - Professional higher technical course 43 - Higher Education - Master (2nd cycle) 44 - Higher Education - Doctorate (3rd cycle) |
| Father's qualification | 1 - Secondary Education - 12th Year of Schooling or Eq. 2 - Higher Education - Bachelor's Degree 3 - Higher Education - Degree 4 - Higher Education - Master's 5 - Higher Education - Doctorate 6 - Frequency of Higher Education 9 - 12th Year of Schooling - Not Completed 10 - 11th Year of Schooling - Not Completed 11 - 7th Year (Old) 12 - Other - 11th Year of Schooling 13 - 2nd year complementary high school course 14 - 10th Year of Schooling 18 - General commerce course 19 - Basic Education 3rd Cycle (9th/10th/11th Year) or Equiv. 20 - Complementary High School Course 22 - Technical-professional course 25 - Complementary High School Course - not concluded 26 - 7th year of schooling 27 - 2nd cycle of the general high school course 29 - 9th Year of Schooling - Not Completed 30 - 8th year of schooling 31 - General Course of Administration and Commerce 33 - Supplementary Accounting and Administration 34 - Unknown 35 - Can't read or write 36 - Can read without having a 4th year of schooling 37 - Basic education 1st cycle (4th/5th year) or equiv. 38 - Basic Education 2nd Cycle (6th/7th/8th Year) or Equiv. 39 - Technological specialization course 40 - Higher education - degree (1st cycle) 41 - Specialized higher studies course 42 - Professional higher technical course 43 - Higher Education - Master (2nd cycle) 44 - Higher Education - Doctorate (3rd cycle) |
| Mother's occupation | 0 - Student 1 - Representatives of the Legislative Power and Executive Bodies, Directors, Directors and Executive Managers 2 - Specialists in Intellectual and Scientific Activities 3 - Intermediate Level Technicians and Professions 4 - Administrative staff 5 - Personal Services, Security and Safety Workers and Sellers 6 - Farmers and Skilled Workers in Agriculture, Fisheries and Forestry 7 - Skilled Workers in Industry, Construction and Craftsmen 8 - Installation and Machine Operators and Assembly Workers 9 - Unskilled Workers 10 - Armed Forces Professions 90 - Other Situation 99 - (blank) 122 - Health professionals 123 - teachers 125 - Specialists in information and communication technologies (ICT) 131 - Intermediate level science and engineering technicians and professions 132 - Technicians and professionals, of intermediate level of health 134 - Intermediate level technicians from legal, social, sports, cultural and similar services 141 - Office workers, secretaries in general and data processing operators 143 - Data, accounting, statistical, financial services and registry-related operators 144 - Other administrative support staff 151 - personal service workers 152 - sellers 153 - Personal care workers and the like 171 - Skilled construction workers and the like, except electricians 173 - Skilled workers in printing, precision instrument manufacturing, jewelers, artisans and the like 175 - Workers in food processing, woodworking, clothing and other industries and crafts 191 - cleaning workers 192 - Unskilled workers in agriculture, animal production, fisheries and forestry 193 - Unskilled workers in extractive industry, construction, manufacturing and transport 194 - Meal preparation assistants |
| Father's occupation | 0 - Student 1 - Representatives of the Legislative Power and Executive Bodies, Directors, Directors and Executive Managers 2 - Specialists in Intellectual and Scientific Activities 3 - Intermediate Level Technicians and Professions 4 - Administrative staff 5 - Personal Services, Security and Safety Workers and Sellers 6 - Farmers and Skilled Workers in Agriculture, Fisheries and Forestry 7 - Skilled Workers in Industry, Construction and Craftsmen 8 - Installation and Machine Operators and Assembly Workers 9 - Unskilled Workers 10 - Armed Forces Professions 90 - Other Situation 99 - (blank) 101 - Armed Forces Officers 102 - Armed Forces Sergeants 103 - Other Armed Forces personnel 112 - Directors of administrative and commercial services 114 - Hotel, catering, trade and other services directors 121 - Specialists in the physical sciences, mathematics, engineering and related techniques 122 - Health professionals 123 - teachers 124 - Specialists in finance, accounting, administrative organization, public and commercial relations 131 - Intermediate level science and engineering technicians and professions 132 - Technicians and professionals, of intermediate level of health 134 - Intermediate level technicians from legal, social, sports, cultural and similar services 135 - Information and communication technology technicians 141 - Office workers, secretaries in general and data processing operators 143 - Data, accounting, statistical, financial services and registry-related operators 144 - Other administrative support staff 151 - personal service workers 152 - sellers 153 - Personal care workers and the like 154 - Protection and security services personnel 161 - Market-oriented farmers and skilled agricultural and animal production workers 163 - Farmers, livestock keepers, fishermen, hunters and gatherers, subsistence 171 - Skilled construction workers and the like, except electricians 172 - Skilled workers in metallurgy, metalworking and similar 174 - Skilled workers in electricity and electronics 175 - Workers in food processing, woodworking, clothing and other industries and crafts 181 - Fixed plant and machine operators 182 - assembly workers 183 - Vehicle drivers and mobile equipment operators 192 - Unskilled workers in agriculture, animal production, fisheries and forestry 193 - Unskilled workers in extractive industry, construction, manufacturing and transport 194 - Meal preparation assistants 195 - Street vendors (except food) and street service providers |
| Admission grade | Admission grade (between 0 and 200) |
| Displaced | 1 – yes 0 – no |
| Educational special needs | 1 – yes 0 – no |
| Debtor | 1 – yes 0 – no |
| Tuition fees up to date | 1 – yes 0 – no |
| Gender | 1 – male 0 – female |
| Scholarship holder | 1 – yes 0 – no |
| Age at enrollment | Age of studend at enrollment |
| International | 1 – yes 0 – no |
| Curricular units 1st sem (credited) | Number of curricular units credited in the 1st semester |
| Curricular units 1st sem (enrolled) | Number of curricular units enrolled in the 1st semester |
| Curricular units 1st sem (evaluations) | Number of evaluations to curricular units in the 1st semester |
| Curricular units 1st sem (approved) | Number of curricular units approved in the 1st semester |
| Curricular units 1st sem (grade) | Grade average in the 1st semester (between 0 and 20) |
| Curricular units 1st sem (without evaluations) | Number of curricular units without evalutions in the 1st semester |
| Curricular units 2nd sem (credited) | Number of curricular units credited in the 2nd semester |
| Curricular units 2nd sem (enrolled) | Number of curricular units enrolled in the 2nd semester |
| Curricular units 2nd sem (evaluations) | Number of evaluations to curricular units in the 2nd semester |
| Curricular units 2nd sem (approved) | Number of curricular units approved in the 2nd semester |
| Curricular units 2nd sem (grade) | Grade average in the 2nd semester (between 0 and 20) |
| Curricular units 2nd sem (without evaluations) | Number of curricular units without evalutions in the 1st semester |
| Unemployment rate | Unemployment rate (%) |
| Inflation rate | Inflation rate (%) |
| GDP | GDP |
| Target | Target. The problem is formulated as a three category classification task (dropout, enrolled, and graduate) at the end of the normal duration of the course |

Conclusiones y modelo



En base a nuestras métricas de desempeño del modelo Gradient Boosting podemos concluir que se puede predecir con hasta un 78% de precisión si un alumno tiene chances de desertar o no.

Para concluir me gustaría utilizar este modelo con datos Chilenos y así ver su real funcionamiento en la sociedad actual.