

The terminology of survival modeling: An insight and alternative modeling of student retention

Dewi Juliah Ratnaningsih
Universitas Terbuka, Indonesia
Asep Saefuddin
IPB University, Indonesia
Anang Kurnia
IPB University, Indonesia
I Wayan Mangku
IPB University, Indonesia

ABSTRACT

Student retention is one indicator of accountability in the implementation of educational programs. Achievement of student retention rates indicates the performance of the quality objectives of an institution or college. To get an accurate picture of the factors related to retention, we need to do modeling. The retention variable is the time response variable measured in semester units. One of the statistical analyzes that can be used to analyze response data in time is survival analysis. The selection of an accurate analytical method in modeling will produce valid conclusions and impact making policies that are right and on target. This paper presents alternative modeling of student retention in distance education using survival analysis. The method used is a literature review. This paper also briefly describes distance education, open and distance education, distance education students' characteristics, distance education student retention, and survival models for modeling student retention in distance education.

Keywords: distance education, open and distance education, student retention, survival analysis.

INTRODUCTION

Data modeling is the process of making a model of a collection of data stored in a database. The purpose of modeling is to visually represent data so that useful information is obtained to policymaking. In general, there is no excellent or perfect modeling produced. It's just that the modeling is expected to be close to the real conditions and useful for stakeholders in making decisions. Data modeling is inseparable from the method or data analysis tool used. Modeling will produce a valid and accurate model and be used to predict when using the right method or analysis. Thus, data modeling is closely related to understanding the selection of an appropriate way or research.

Student retention is one indicator of institutional accountability in the implementation of educational programs. A low student retention rate indicates that the quality of the institution's goals is good. The institution has a graduate quality that can be relied upon and accepted by the community. Therefore, retention becomes essential for a tertiary institution in determining the quality of education services. Likewise, with the Universitas Terbuka (UT) Indonesia implementing the Open and Distance Education (ODE) system, student retention is an indicator of education quality assurance.

Student retention conditions need to be appropriately modeled to obtain precise and accurate information following the data characteristics of existing data. The use of appropriate and appropriate analytical methods can produce valid conclusions. For this reason, a good understanding of the proper statistical way or analysis and following the characteristics of student retention data needs to be done.

This paper explains insight into modeling student retention at the ODE using a survival model by considering existing data.

SELECTED LITERATURE REVIEW

Distance Education

History of Distance Education (DE) begins in the 1700s and 1800s, reflecting egalitarianism in the world of education (Casey 2008). In brief, DE's historical development can be grouped based on the dominant technology it uses, as shown in Figure 1 (Bozkurt 2019).

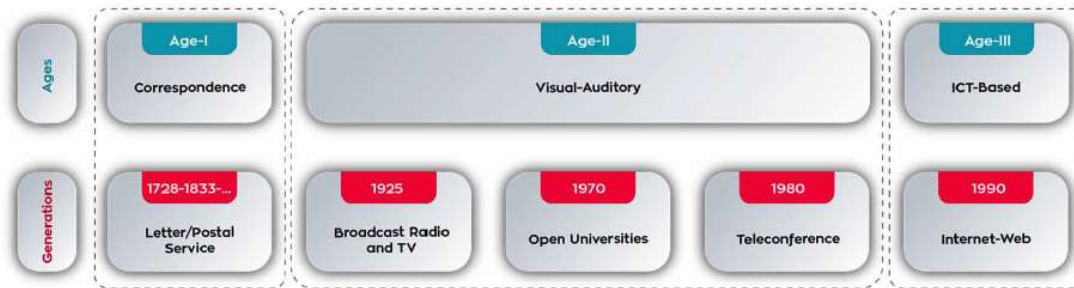


Figure 1. History of the development of distance education.

Meanwhile, Taylor (2001) classifies of DE generation into five (5) generations, namely: (1) correspondence models, (2) multimedia models, (3) tele-learning models, (4) flexible learning models, and (5) a smarter flexible learning model (The Intelligent Flexible Learning Model). The term DE has several alternative names as stated by Bozkurt (2019), including open learning, open teaching, non-traditional education, distance learning, distance teaching, correspondence education, independent study, flexible education, flexible learning, home study, life-long education, lifelong learning.

UNESCO (2002) defines DE as an educational process where all or most of the teaching is carried out by someone who is transferred in the space and time of the student, with the effect that all or most of the communication between the teacher and students is through artificial media, either electronic or print. Schlosser and Simonson (2009) define DE as an institution-based education system. Learning groups are separated, and where interactive telecommunications systems are used to connect learners, resources, and instructors. Bates (2005) defines DE as an education system where there is a separation of students in space and time from other students, learning resources, and teachers.

Moore and Kearsley (2011) define DE as planned teaching and learning. Education usually takes place in a different place than knowledge, which requires communication through technology and specialized institutional organizations. Portway and Lane (1997) define DE as a teaching and learning system, where instructors and students are geographically separated. Therefore the role of electronic devices and print media and technology is significant in delivering learning materials. Mehrotra et al. (2001) state that DE is a formal approach to instruction, where most teaching takes place while educators and students are not in the same physical condition. Picciano (2001) states that DE is an education system that uses three forms of media today: television, two-way video conferencing, and asynchronous learning networks. Furthermore, Birnbaum (2001) suggests that asynchronous DE provides multi-modal, web-based delivery for instructions that students can review anytime and anywhere. This condition allows students to access material, lectures, and education anywhere and anytime, depending on student readiness.

Distance education is considered a more democratic form of education because it aims to reach all strata/sections of society (Gunawardena & McIsaac, 2004). Simonson et al. (2015) state that DE provides more extensive access to learning for people who do not have the opportunity to learn in face-to-face lectures. Mehrotra et al. (2001), "distance learning, or distance education, is not the future of the possibility that tertiary education must prepare, it is the present reality of creating opportunities and challenges for educational institutions; the fact that offers students expanded choices on where, when,

how and from whom they study; making the reality of education accessible to more and more people. Although, in practice, DE has been criticized as having lower quality and effect compared to campus-based schools, some studies suggest that the DE model is more effective than the traditional campus-based education model.

From the various definitions that have been stated previously, then in DE required qualified technology to bridge the separation of students, instructors, and learning resources. Meyer (2002) says that to help alleviate the demands of separation between teachers and students, institutions have begun to utilize available technologies, such as audio connections (i.e., telephones), video cassettes, and television. Starting in the 1980s, popular media in DE used communication satellites to send broadcasts of lectures and instructions to students far from campus locations. In the late 1980s to 1990s, the use of interactive video-based media.

In detail, Taylor (2001) summarizes the five generations of technology concerning the flexibility and ability to facilitate interactions in learning and the cost aspects, as presented in Table 1.

Table 1. Technology-Based on Characteristics in Distance Education

Models of Distance Education and Delivery Technologies	Characteristics of Delivery Technologies					
	Flexibility			Highly Refined Materials	Advanced Interactive Delivery	Institutional Variable Costs Approaching Zero
	Time	Place	Pace			
FIRST GENERATION						
The Correspondence Model						
• Print	Yes	Yes	Yes	Yes	No	No
SECOND GENERATION						
The Multimedia Model						
• Print	Yes	Yes	Yes	Yes	No	No
• Audiotape	Yes	Yes	Yes	Yes	No	No
• Videotape	Yes	Yes	Yes	Yes	No	No
• Computer-based Learning	Yes	Yes	Yes	Yes	Yes	No
• (e.g. CML/CAL/IMM)						
• Interactive video (disk and tape)	Yes	Yes	Yes	Yes	Yes	No
THIRD GENERATION						
The Telelearning Model						
• Audio tele-conferencing	No	No	No	No	Yes	No
• Video-conferencing	No	No	No	No	Yes	No
• Audiographic communication	No	No	No	Yes	Yes	No
• Broadcast TV/Radio and audio-teleconferencing	No	No	No	Yes	Yes	No
FOURTH GENERATION						
The Flexible Learning Model						
• Interactive multimedia (IMM) online	Yes	Yes	Yes	Yes	Yes	Yes
• Internet-based access to WWW resources	Yes	Yes	Yes	Yes	Yes	Yes
• Computer-mediated communication	Yes	Yes	Yes	Yes	Yes	No
FIFTH GENERATION						
The Intelligent Flexible Learning Model						
• Interactive multimedia (IMM) online	Yes	Yes	Yes	Yes	Yes	Yes
• Internet-based access to WWW resources	Yes	Yes	Yes	Yes	Yes	Yes
• Computer-mediated communication, using automated response systems	Yes	Yes	Yes	Yes	Yes	Yes
• Campus portal access to institutional processes and resources	Yes	Yes	Yes	Yes	Yes	Yes

Open and Distance Education

The Open and Distance Education (ODE) System comes from two terms: the open education system and the distance education (DE) system. The concept of open education (open education or open learning) is a goal or ideals of policy regarding the education system. Emphasizing the importance of system flexibility to minimize place, time, and aspect constraints caused by student characteristics such as economic conditions (Bates, 1995).

The term open in the open education system means more free from limitations and has absolutely nothing to do with the ODE system. The ODE system refers more to the system or mode of delivery of the learning process. The open education system relates to changes in educational organizations' structure, becoming an open organization in place, time, learning material, learning system, etc. Open education is a mindset and approach used to provide a variety of choices in learning for students and give as much control as possible for students to determine what will be discovered and learning strategies.

Belawati (2019) states that not all distance learning is open. Some literature indicates that the characteristics of open education must at least contain an element of flexibility, including aspects of age (no age restrictions), location (bias from anywhere), costs (cheap or even free), length of study (no limit on study time), and prerequisites (no need to have a previous education diploma), multi-entry and multi-exit (can enter and stop at various time / anytime alternatives). Furthermore, Belawati (2019) described DE and ODE as an incision depicted in Figure 2.

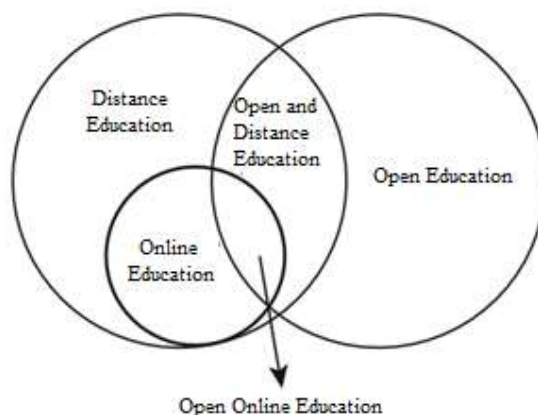


Figure 2. Distance Education, Online Education, Open and Distance Education

Bozkurt (2019) emphasizes that the term ODE is more widely used to reflect a shift toward a more social and learner-centered view of learning that adopts openness to broader social justice. ODE is an essential element in the education and training system in the future. Currently, ODE plays a crucial role in determining the creation of a global knowledge-based society.

UNESCO (2012) reaffirmed the ODE's role in diversifying the education delivery system, particularly for technical and vocational education, encouraging cooperation and partnerships between companies, professional bodies, and distance teaching institutions. Support is also given to ODE to meet the unique needs of disabled people, migrants, cultural and linguistic minorities, refugees, populations in crises, which cannot be efficiently achieved by traditional shipping systems. Besides, the ODE is deeply embedded in teacher education, correctly, for the training of teachers in positions and the practice of teacher educators.

One example of the "open" movement in education is reflected in OER (Open Education Resources) and MOOCs (Massive Open Online Courses), which is an example of new developments using a range of connected technologies. The main idea of the emergence of the OER movement is the idea of the world of knowledge, namely public goods that technology in general and the worldwide web (www) provides opportunities for everyone to share, use and reuse. In other words, the OER movement is also driven by social responsibility, which aims to provide fair and universal access to knowledge and web platforms, in particular, to function as distribution platforms. This driving force behind the OER movement has encouraged various initiatives around the world to provide access to various educational resources, including lecture material as well as educational material.

Two official definitions are available for OER; one from UNESCO and one from OECD. UNESCO (2002) defines OER as digital teaching material for the learning and research process. This OER is packaged in various digital or other media, which are in the public domain or have been released under an open license, which can be accessed without charge, and can be distributed to others without any restrictions. Meanwhile, UNESCO (2012) explained that OER is a digital material offered freely and openly for educators, students, and independent students and is used again for teaching, learning, and research. Based on this definition, OER emphasizes more on "openness," "digital" format, and "reusability/adaptation" or resources.

Besides OER, other open movements are MOOCs. MOOCs are referred to as "the great revolution in education" (Bates 2014). Simoson et al. (2015) stated that digital technology, OER, and MOOCs provide more extensive access to learning for people who do not have the opportunity to learn in face-to-face lectures. Through OER and MOOCs, the initial goals of distance education, social justice and

openness for all, and reflecting community-based and socially driven learning approaches will soon be realized. Bates (2014) further identified MOOCs design practices, namely

1. use of social media (online courses supported by various 'connected' tools and media),
2. material content is determined by the participants or students who decide and contributing material content,
3. well-distributed communication (communication is an automatic network with many sub-components), and
4. students decide for themselves whether what they have learned is appropriate for them.

Anderson and Simpson (2012) suggest that cellular technology helps in the development of ODE. Mobile technology can improve learning anytime and anywhere without the constraints of space and time and distance.

Characteristics of Distance Education Students

The characteristics of DE students are unique and diverse (Cercione 2008). The tendency of students at DE institutions is adult students, which varies depending on age, sex, knowledge, and skills, and the context of the place of learning (Kara et al. 2019). Knowles (1990) suggests that one of the hallmarks of adult learning is that they are motivated to learn if they have perceived needs and interests that can be fulfilled through education; therefore, the starting point for learning is their needs and interests. Merriam and Caffarella (1999) state that most adult students are highly motivated and task-oriented. Furthermore, Merriam and Caffarella (1999) state that adult students have responsibilities such as family, work, and situations (e.g., childcare, need for income) that can disrupt the learning process. Tladi (2013) states that 62.8% of DE students are not ready for the exam because they do not have enough time to study. Furthermore, Tladi (2013) revealed that DE students are generally married, have families and responsibilities, so they lack the commitment to explore independently.

Schuemmer (1993) suggested that in the DE system, students' learning process is more complicated because, in general, DE students are elderly, work, and have a family. Some researchers also stated the same thing. Age factors are very influential in the DE system (Valasek, 2001). Besides, gender factors (Aragon and Johnson 2008), race/ethnicity (Moore et al. 2002, Sullivan 2001), financial needs (Parker 2003), and GPA (Holder 2007, Harrel and Bower 2011). The variety of conditions and characteristics of UT students causes them to be required to be able to coordinate various aspects, such as family, work, and free time with study time. Given these conditions, their motivation to attend lectures in the DE system is also very diverse.

Regarding students' motivation to join DE institutions, some students stated that they joined DE on their initiative or at the institution's instigation/assignment where they worked. Many DE students end up dropping out. Motivation/psychology plays an essential role in the perseverance of DE students to survive to complete their education. Personal goals, a sense of community, and family support are also psychological motivators that increase student persistence (Yang et al. 2017).

Student Retention of Distance Education

The term retention has various meanings. Some researchers have conducted studies regarding the retention of DE students, including Belawati (2006) and Arifin (2018). Belawati (1998) defines retention as the completion of courses and continuing re-registration. Student retention is expressed as a commitment to completing studies within the given timeframe and re-registering to continue education and maintain their educational status without significant disruption. Sembiring (2014) states that student retention is conceptually defined as a commitment to consistently reuse favored services in the future, namely to re-register. Meanwhile, Arifin (2018) says that student retention is related to student behavior, continuing learning in a particular program by registering for four consecutive semesters.

Low student retention is the biggest problem and has a significant contribution to student studies' failure in the DE program (Rovai and Downey 2010). From an institutional perspective, retention serves as a measure of institutional accountability in the effectiveness and quality of the implementation of educational programs (Tinto 2006). Besides, Kim et al. (2010) state that retention shows the

achievement of an institution's mission to educate and prepare students to be able to survive to complete their studies. Dropping out of school indicates the failure to develop and support students and is a significant loss for an institution. Belawati (1998) states that organizationally, low retention can endanger UT's existence; a decrease in the number of students means a reduction in UT's financial income. In turn, this will affect the institution's ability to expand and improve the quality of teaching and education services. A low retention rate can jeopardize an institution's image because it appears to assist with a flawed system.

Several studies conducted by DE experts show that student retention in DE is low, lower than conventional/face-to-face education. Xu and Jaggars (2011) stated that DE's student retention was more depressed by around 8-14% compared to face-to-face learning. Moore and Fetzner (2009) argue that student retention rates in distance classes are claimed to be 10-20% lower than conventional education. Likewise, Simpson's (2003) study in England stated the same thing. The percentage of remote institutional retention is more moderate than in traditional institutions (Figure 2.4). Belawati (1998) has conducted a UT study, and she stated that the average retention rate of UT students whose first registration between 1984 and 1990 was only 4.8%.

Retention is a complex and complicated problem. Various factors influence student retention at DE. Choi (2017) provides a concept regarding students' dropout rate at DE, as presented in Figure 2.5. Figure 2.5 shows that retention is influenced by various factors, namely: learner factors, external factors, internal factors, and external factors. Pierrakeas et al. (2004) stated that younger students (<30 years) had low retention. This fact is because they do not yet have an independent learning experience, and they tend to underestimate the effort and workload needed for study at the university level. Xenos et al. (2002) and Pierrakeas et al. (2004) in Greece stated that there is a correlation between age and dropping out of DE students.

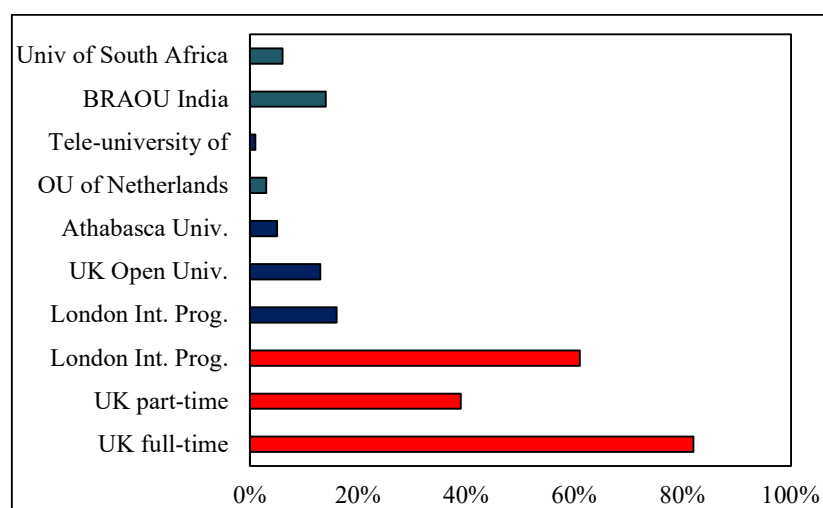


Figure 3. Percentage of student retention in higher education institutions

Rovai (2008) states, in general, the factors that cause dropouts experienced by DE students include old age, lack of study time, difficulties in accessing the internet, lack of feedback from tutors, work, family, external stimuli, and personal financial problems.

Besides, several studies have shown that GPA is very influential in student retention and is a determining factor for universities' sustainability at universities (McCormic and Lucas 2014). Students' academic characteristics are determinants of dropout students (Ratnaningsih et al. 2008; Boton and Gregory 2015). Another factor is the credit load taken by students. Cambuzzi et al. (2015) in Brazil stated that many students dropped out of college because the credit load did not match students' abilities. Allen et al. (2016) said that many students experienced dropouts due to taking too much course material and paying tuition fees.

Muljana and Luo (2019) made a fairly detailed and comprehensive summary of the factors influencing student retention in DE from several research results. These factors are seen from various

aspects, both institutions, instructors, and students. Institutional factors that influence student retention include institutional encouragement in student services and orientation, tutoring services, technology support, institutional understanding of student conditions and needs, curriculum or program difficulty level, and logical course structure. Boston et al. (2011) in Muljana and Luo (2019) suggested that a too easy or too tricky curriculum can affect student retention.

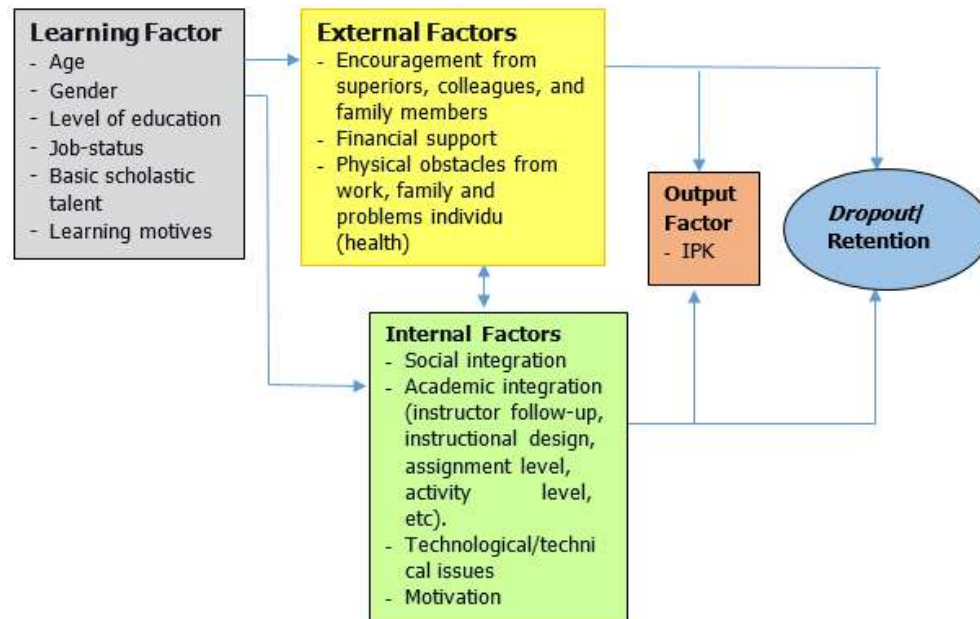


Figure 4. Conceptual model of failure rate in Distance Education

Instructor/instructor factors that influence student retention include active interaction between instructors and students, social interaction between students, so they do not feel isolated, student activeness in discussions, tutors activeness in learning activities, providing guidance and feedback by tutors on discussions or student assignments, effective communication between tutors and students, attractive learning designs, interactive learning materials, clear instructions from the instructor, and the ease of obtaining teaching materials. From the student's perspective, factors that influence student retention include behavioral characteristics, demographic characteristics, and other personal characteristics. Student behavior attributes include self-regulation, metacognition, self-efficacy, discipline, motivation, learning strategies, learning satisfaction, learning readiness, technological skills, setting clear learning goals, and time management. Demographic characteristics include age and gender (although gender is not always associated with student retention in DE). Wladis et al. (2015) in Muljana and Luo (2019) suggested that older students tend to perform better and tend to have high retention in learning. Meanwhile, other personal characteristics that influence student retention include family support, home environment, family responsibilities, work, responsibilities and workload, financial problems, life problems (related to health), class conditions, and GPA.

Coffield et al. (2004) suggested that learning in the DE system can be improved if students are motivated to learn by understanding their learning styles. Merriam and Caffarella (1999) report that education is a process, so there is no successful adult learning theory applied to all adult learning environments. This fact is supported by Cercone (2008), which states that adult learners in DE are diverse and have their characteristics. Every individual is different; learners at DE are unique and have different needs. This condition shows that each individual is not the same, and learning motivation will undoubtedly be extra. Thus, each individual's risk to be able to survive well in lectures will indeed be different. Also, student retention cannot be separated or bound by socio-cultural factors or the social culture of a country (Belawati 1998, Rovai 2008, and Holder 2007). Diversity in educational and organizational culture, geography, technology, study programs, student characteristics, development of

support systems from institutions all play a significant role in increasing student retention in the context of DE.

Survival Model for Modeling Students Retention of Distance Education

One indicator of institutional accountability in organizing educational programs is student retention. Low student retention indicates an inability of institutions to improve the quality of education services. The issue of student retention needs to be the main focus of education services, including the ODE. Information related to factors influencing student retention needs to be explored carefully, precisely, and accurately. This information can be used to consider institutions in policymaking. With adequate, accurate, and accurate student retention modeling, ODE organizers are expected to increase student retention in completing studies and maintain student Rough Participation Rates (GER). Student retention and GER are indicators of the quality and success of Higher Education in Indonesia.

In statistics, the analytical method used for modeling survival time data is survival analysis. Survival analysis is a set of statistical procedures used to analyze data where the variable of concern is until the event occurs (Kleinbaum and Klein 2012). One of the aims of survival analysis is to discover the connection between event time and predictor variables (covariates), which are possibly significant for the process. The time until a particular event occurs is called a survival time/failure time. Survival time can be stated in years, months, weeks, or days from the beginning of the study until an event occurs.

In survival analysis, response variables related to time can be grouped as complete data or incomplete data. Comprehensive data is data when the time of occurrence of events can be observed during the study period. Meanwhile, insufficient information is data when the phenomenon of events cannot be observed ultimately. Incomplete data in survival analysis is often called censored data. D. R Cox, in 1972, developed a regression analysis called Cox regression or better known as the Cox proportional hazard model.

Some researchers have conducted studies on student retention in DE. However, the research has not reviewed that student retention problems can be seen as problems related to event analysis time or time. Low student retention can lead to dropping out of college (failure to complete study). Failure to conduct an investigation can be seen as a failure time. Thus, modeling of student retention can be seen as modeling survival time. Therefore, analysis of modeling data retention of students can use survival analysis. An illustration that student retention in the ODE can be seen as a survival time can be illustrated in 7 cases that may occur while attending lectures.

Figure 5 shows the various possible causes of student academic travel during lectures at the ODE. In the first case, students smoothly attend classes until finally they are declared graduated. They can go on an educational journey without a time lag to stop temporarily. They attend lectures according to the allotted time. In the second case, students experience academic leave several semesters. They were declared non-active because they did not register for four consecutive semesters. The conditions of their educational trips are on / off. Until some time has been determined to be non-active status.

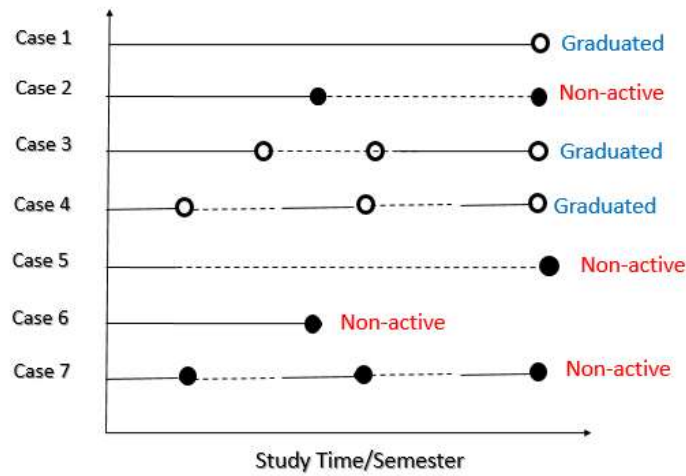


Figure 5. Cases of Student Academic Track during Lecture at Open Distance Education

Students can complete their studies in the third and fourth cases even though they have on/off trips. But not until experiencing non-active. This condition is possible because, in general, ODE students already work and have families. In the fifth case, students only re-register / register courses at the beginning of the semester. The rest, they do not register terms until the status becomes non-active. In the sixth case, students from the beginning did not re-register. They only register initially and do not continue to enroll in semester one and the following semester. Until a few semesters, they did not register for the course. In the seventh case, students experienced significant fluctuations. During their lectures, they experienced non-active status several times. In the end, they remained non-active.

The academic track conditions, illustrated in Figure 5, often occur and are real events that appear at the Open University, Indonesia. Open University is the only Higher Education institution that implements the ODE system. Based on the data conditions depicted in Figure 5, the UT student retention can be viewed as a matter related to time or time analysis. Therefore, alternative modeling for student retention cases can use survival analysis.

The factors that influence student retention are very complex and varied. In general, these factors can be grouped into individual/learner factors, internal factors, and external factors. Student retention is very personally tied to the student himself, where each individual has different characteristics and learning needs. Besides, retention is also related to the socio-cultural factor or social culture of a country. Diversity in educational and organizational culture, geography, technology, study programs, student characteristics, development of institutional support systems all play a significant role in increasing the retention of DE students.

In modeling, influential factors are called covariates. There are very many covariates involved in modeling student retention. These covariates can be grouped into time-independent covariates, time-dependent covariates, and random covariates. Covariates are time-independent, including gender, marital / family status, employment status, education level / educational background, student residence, and curriculum structure. The covariates depend on time, among others: age, credit load is taken, many courses registered, GPA, and acquisition of teaching materials (Ratnaningsih et al. 2019)

Meanwhile, covariates included random effects observed have study programs of interest, tutoring services, tutorials, student activeness in discussions, tutorial assignments, attendance of tutors, feedback by tutors in meetings or tasks, and communication effective between tutors and students. Covariates of unobserved random effects include motivation, learning time management, institutional support in technology, student social interaction, student learning styles, institutional understanding of student needs, and organizational culture within the institution. With the diversity of these covariates, the survival analysis modeling using the Cox model is inadequate. Modeling that can be done is incomparable risk modeling by considering various existing covariates. Therefore, in data retention, possible data analysis is a survival model with a non-proportional hazard.

The roadmap for survival analysis and its differences with ordinary regression analysis and basic terms in survival analysis is presented in Figure 6. From Figure 6, it can be explained that if the response variable is censored data, the use of regression analysis will produce bias in estimating standard parameters and errors (Henderson and Oman 1999). Therefore, if the response variable contains censored data, then the use of survival analysis is more appropriate. Here is a definition of some general terms contained in survival analysis that need to be known (Ratnaningsih et al. 2019).

1. Censored data is one form of incomplete observation. It means that it did not experience any events during the study period. In this study, censored data is data of active or graduated students during the study period.
2. Events are events that are observed in individuals during the study period.
3. Survival time is the time of resilience, i.e. when a subject/individual goes through something from a specific time until an event occurs.
4. Failure Time is the time of "failure" experienced by the subject/individual during the study period.
5. Time to an event is the time experienced by the subject/individual until an event occurs.
6. Lifetime is the lifetime of the subject observed in survival analysis.
7. Cox proportional hazard is a Cox model with relative risks that the level of risk for one individual is proportional to the level of risk of another individual in a constant ratio over time.
8. Non-Proportional Hazard means that the subject/individual observed has different risks and can change according to time changes.

9. The hazard ratio is the level of risk or danger of an individual to the observed covariate.
10. Covariates are factors in modeling that act as explanatory variables.
11. Time independent covariates are covariates (factors that act as explanatory variables) in modeling that are not time-dependent—for example, gender, domicile, educational background.
12. Time-dependent covariates are covariates (factors that act as explanatory variables) in time-dependent modeling. For example: taking courses, the number of credits that can change each semester.
13. The fixed effect is a permanent effect on a model that functions as a covariate (explanatory variable).
14. Random effects are random effects that can be involved in modeling.
15. Frailty is a proportionality factor that is not observed random effects that modify individuals or related individuals' hazard function, for example, in cultural health/hospital culture in treating patients, patient genetics, or patient dietary patterns. For instance, in education, student learning culture, environmental factors, or student learning time management.

The use of survival models in modeling student retention can produce more valid and accurate analysis results. In modeling, it considers censored data that is data of students who are on / off in registering courses per semester. This condition is thought to influence the resulting modeling. To use appropriate statistical analysis to produce valid and accurate conclusions, it is necessary to understand the data structure and the assumptions underlying the research's use. The study of UT student retention can be read in studies conducted by Ratnaningsih et al. (2008), Ratnaningsih et al. (2018), and Ratnaningsih et al. (2019).

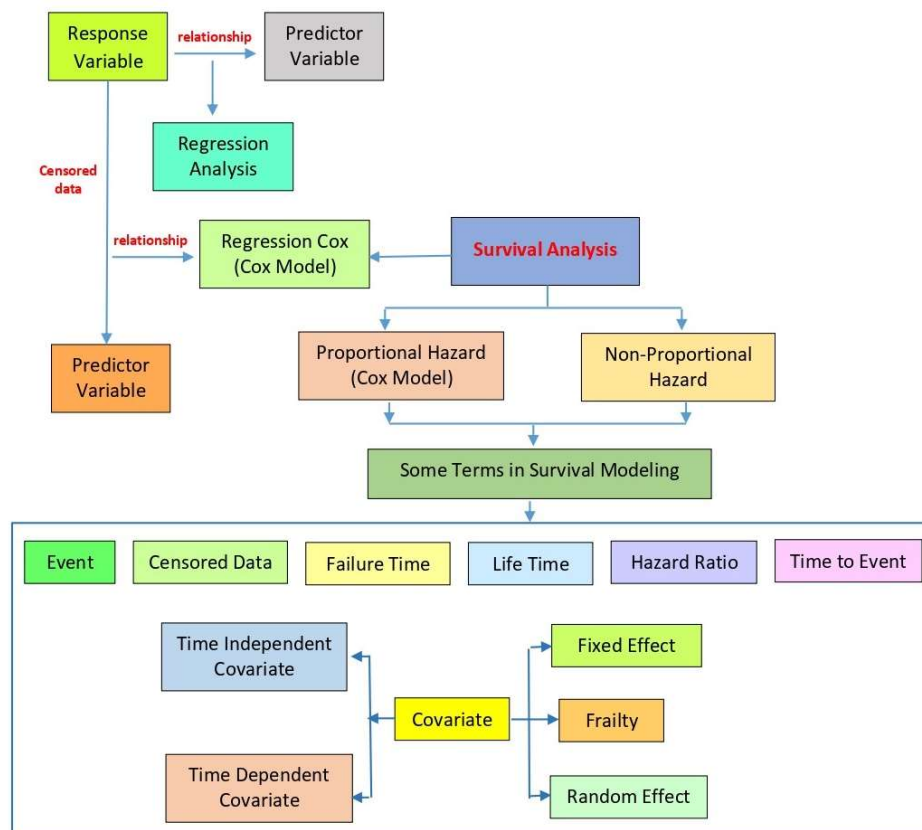


Figure 6. Roadmap for Survival Analysis

CONCLUSION

Modeling is something that needs to be done to simplify existing problems. Good modeling can produce a valid and accurate outcome. The accuracy of using modeling methods requires a good understanding of the data and assumptions that underlie the use of the method or analysis tool used.

Student retention is one indicator of an education provider's accountability, including Distance Education and Open and Distance Education. The factors that influence retention are very complex. Retention problems can be viewed as problems that are related to time or time analysis of events. Statistical analysis to analyze time events is survival analysis. Survival analysis used to analyze the relationship between factors or covariates is the Cox analysis or the Cox model.

In the case of data retention, that is influenced by several factors, namely: individual/learner factors, internal factors, and external factors. These factors can be categorized as time-dependent covariates, time-dependent covariates, and random effects. The use of Cox regression or Cox models for data retention modeling involving various factors is inadequate. Survival models that can overcome multiple covariates are survival models with incomparable risk by considering the different types of covariates engaged in modeling.

REFERENCES

- Anderson, B., & Simpson, M. (2012). History and heritage in distance education. *Journal of Open, Flexible and Distance Learning*, 16(2), 1-10. Retrieved from <http://files.eric.ed.gov/fulltext/EJ1080085.pdf>.
- Allen, I. E., Seaman, J., Poulin, R., & Straut, T. T. (2016). *Online report card: Tracking online education in the United States*. Babson Park, MA: Babson Survey Research Group Retrieved from <http://onlinelearningsurvey.com/re-ports/online-report-card>.
- Aragon, S. R., & Johnson, E. S. (2008). Factors influencing completion and noncompletion of community college online courses. *The American Journal of Distance Education*, 22(3), 146–158.
- Arifin, H. (2018). The role of student support services in enhancing student persistence in the open university context: lesson from Indonesia Open University. *Turkish Online Journal of Distance Education-TOJDE*, 19(3), 156-168.
- Bates, A.W. (1995). *Technology, open learning, and distance education*. New York, N.Y.: Routledge.
- Bates, T. (2005). *Technology, e-learning and distance education*. New York: Routledge.
- Bates, T. (2014). *Comparing xMOOCs and cMOOCs: Philosophy and practice*. Retrieved 04 01, 2016, from Online learning and distance education resources: <http://www.tonybates.ca/2014/10/13/comparing-xmoocs-and-cmoocs-philosophy-and-practice/>.
- Belawati, T. (1998). Increasing student persistence in Indonesian post-econdary distance education. *Distance Education*, 19(1), 81-108. DOI: 10.1080/-0158791980190107.
- Belawati, T. (2006). The impact of online tutorials on course completion rates and student achievement. *Learning, Media and Technology*, 30(1), 15-25. DOI:10.1080/13-581650500075520.
- Belawati, T. 2019. Online learning. Kementrian Riset, Tenlogi, dan Pendidikan Tinggi. Tangerang Selatan: Universitas Terbuka.
- Birnbaum, B.W. (2001). Foundations and practices in the use of distance education. *Mellen Studies in Education*, 66, 1-174.
- Boston, W. E., & Ice, P., & Gibson, A. M. (2011). Comprehensive assessment of student retention in online learning environments. *Online Journal of Distance Learning Administration*, 14(1), 1-19.
- Boton, E. C., & Gregory, S. (2015). Minimizing attrition in online degree courses. *Journal of Educators Online*, 12(1), 62-90.
- Bozkurt, A. (2019). From Distance Education to Open and Distance Learning: A Holistic Evaluation of History, Definitions, and Theories. In S. Sisman-Ugur, & G. Kurubacak (Eds.), *Handbook of Research on Learning in the Age of Transhumanism* (pp. 252-273). Hershey, PA: IGI Global.

- Cambruzzi, W., Rigo, S. J., & Barbosa, J. L. V. (2015). Dropout prediction and reduction in distance education courses with the learning analytics multitrait approach. *Journal of Universal Computer Science*, 21(1), 23-47. DOI: [10.3217/jucs-021-01-0023](https://doi.org/10.3217/jucs-021-01-0023).
- Casey, D. M. (2008). The Historical Development of Distance Education through Technology. *TechTrends*, 52(2), 45–51.
- Cercone, K. (2008). Characteristics of adult learners with implication for online learning design. *AACE Journal*, 16(2), 137-159.
- Choi, H. J., & Kim, B. U. (2017). Factors affecting adult student dropout rates in the Korean Cyber-University degree programs. *The Journal of Continuing Higher Education*, 00, 1–12. DOI: 10.1080/07377363.2017.1400357.
- Coffield, F., Moseley, D., Hall, E., & Ecclestone, K. (2004). *Learning Styles and Pedagogy in Post-16 Learning: A Systematic and Critical Review*. LSRC reference. Learning & Skills Research Centre. London. Retrieved on September 15, 2019 from www.voced.edu.au/content/ngv%3A13692.
- Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(2), 187-220.
- Gunawardena, C. N., & McIsaac, M. S. (2004). Distance education. In D. H. Jonassen, *Handbook of Research in Educational Communications and Technology* (2nd ed ed., pp. 355-395). Mahwah, NJ: Lawrence Erlbaum Associates.
- Harrell, I. L., & Bower, B. L. (2011). Student characteristics that predict persistence in community college online courses. *American Journal of Distance Education*, 25(3), 178-91.
- Holder, B. (2007). An investigation of hope, academics, environment, and motivation as predictors of persistence in higher education online programs. *Internet & Higher Education*, 10(4), 245–260.
- Kara, M., Erdogdu, F., Kokoc, M., & Cagiltay, K. (2019). Challenges faced by adult learners in online distance education: a literature review. *Open Praxis International Council for Open and Distance Education*, 11(1), 5-22.
- Kim, E., Newton, F., Downey, R., & Benton, S. (2010). Personal factors impacting college student success: Constructing college learning effectiveness inventory (CLEI). *College Student Journal*, 44(1), 112-125. DOI: 1984269161.
- Kleinbaum, D. G., & Klein, M. (2012). *Survival Analysis: A Self-Learning Text*. Ed ke-3. Gail M, Krickeberg K, Samet JM, Tsiatis A, Wong W, editor. New York (US): Springer. DOI: 10.1007/978-1-4419-6646-9.
- Knowles, M. S. (1990). *The Adult Learner: A Neglected Species* (4e). Houston: Gulf Publishing.
- McCormick, N. J., & Lucas, M. S. (2014). Student retention and success: Faculty initiatives at Middle Tennessee State University. *Journal of Student Success and Retention*, 1(1), 1-12.
- Mehrotra, C.M., Hollister, C.D., & McGahey, L. (2001). *Distance learning: Principles for effective design, delivery, and evaluation*. Thousand Oaks, CA: Sage Publications, Inc.
- Merriam, S. B., & Caffarella, R. S. (1999). *Learning in Adulthood* (2nd ed.). San Francisco: Jossey-Bass.
- Meyer, K. A. (2002). Quality in Distance Education: Focus on On-line Learning. In A.J. Kezar (Ed.), *ASHE-ERIC Higher Education Report*, 29, 1-134.
- Moore, K., Bartkovich, J., Fetzner, M., & Sherrill, I. (2002). Success in Cyberspace: Retention in Online Courses. Paper was presented at the Association for Institutional Research. Toronto: Canada.
- Moore, J. C., & Fetzner, M. J. (2009). The road to retention: A closer look at institutions that achieve high course completion rates. *Journal of Asynchronous Learning Networks*, 13(3), 1-22. DOI: 10.24059/olj.v13i3.1650.
- Moore, M. G., Kearsley, G. (2011). *Distance education: A systems view of online learning* (3rd ed.). Belmont, CA: Wadsworth Cengage Learning.
- Muljana, P. S., & Luo, T. (2019). Factors contributing to student retention in online learning and recommended strategies for improvement: A systematic literature review. *Journal of Information Technology Education*, 18, 19-57.
- Parker, A. (2003). Identifying predictors of academic persistence in distance education, *USDLA Journal*. 17(1), 55–62.
- Picciano, A. (2001). *Distance learning: Making connections across virtual space and time*. Upper Saddle River, NJ: Merrill Prentice Hall.

- Pierrakeas, C., Xenos, M., Panagiotakopoulos, C., & Vergidis, D. (2004). A comparative study of dropout rates and causes for two different distance education courses. *International Review of Research in Open and Distance Learning*, 5(2), 1-15.
- Portway, P. S., & Lane, C. (1997). Guide to teleconferencing and distance learning (3 rd ed.). Livermore, CA: Applied Business Telecommunications.
- Ratnaningsih, D. J., Saefuddin, A., & Wijayanto, H. (2008). An analysis of drop-out students' survival in distance higher education. *Journal of Open Distance Education*, 9(2), 101–110.
- Ratnaningsih, D. J., Saefuddin, A., Kurnia, A., & Mangku, I. W. (2019). Stratified-Extended Cox Model in Survival Modeling of Non-Proportional Hazard. *IOP Conference Series: Earth and Environmental Science*, 299(012023), 1-18. DOI: 10.1088/1755-1315/299/1/012023.
- Ratnaningsih, D. J., Saefuddin, A., Kurnia, A., & Mangku, I. W. (2019). Mixed Models of Non-Proportional Hazard and Their Applications in the Distance Education Students Retention Data. *Dissertation*. Departement of Statistics. IPB University: Bogor.
- Ratnaningsih, D. J., Saefuddin, A., Kurnia, A., & Mangku, I. W. (2020). Stratified-extended Cox with frailty model for non-proportional hazard: A statistical approach to student retention data from Universitas Terbuka in Indonesia. The paper was accepted in *Thailand Statistician* and will be published as soon as possible.
- Rovai, A. P. (2008). In search of higher persistence rates in distance education online programs. *The Internet and Higher Education*, 6(1), 1-16. DOI:10.1016/S1096-7516(02)00158-6.
- Rovai, A. P., & Downey, J. R. (2010). Why some distance education programs fail while others succeed in a global environment. *Internet and Higher Education*, 13, 141-147. DOI:10.1016/j.iheduc.2009.07.001
- Schlosser, L. A., & Simonson, M. R. (2009). *Distance Education: Definitions and Glossary of Terms* (2nd ed.). Charlotte, NC: Information Age Publishing.
- Schuemer, R. (1993). *Some Psychological Aspects of Distance Education*. Hagen, Germany: Institute for Research into Distance Education. (ED 357 266.).
- Sembiring, M. G. (2014). Modeling the determinants of student retention in distance education institutions. *International Journal of Continuing Education and Lifelong Learning*, 6(2), 15-28.
- Simonson, M., Smaldino, S., & Zvacek, S. (2015). *Teaching and learning at a distance foundations of distance education*. Charlotte, North Carolina: Information Age Publishing, Inc.
- Simpson, O. (2013). Student retention in distance education: are we failing our students? *Open Learning: The Journal of Open, Distance and E-Learning*, 28(2), 105–119. DOI: 10.1080/02680513.2013.847363.
- Sullivan, P. (2001). Gender differences and the online classroom: Male and female college students evaluate their experiences. *Community College Journal of Research and Practice*, 25, 805–818.
- Taylor, J. C. (2001). *Fifth generation distance education*. Higher Education Division, Department of Education, Training and Youth Affairs.
- Taylor, J. (2000). New millennium distance education. Dalam V. Reddy & S. Manjulika (Eds). *The world of open and distance learning*. New Delhi:Viva. Retrieved from www.usq.edu.au/users/taylorj/publications_presentations/2000IGNOU.doc.
- Tladi, L. S. (2013). Factors affecting examination attrition: does academic support help? A survey of ACN203S (Cost Accounting and Control) students at Unisa. *Open Learning: The Journal of Open Distance and e-Learning*, 28(1), 67-82. DOI: 10.1080/02680513.2013.782812.
- Tinto, V. (2006). Research and practice of student retention: What next? *Journal College Student Retention*, 8(1), 1-19. DOI: 10.2190/4YNU-4TMB-22DJ-AN4W.
- UNESCO. (2002). *Open and Distance Learning: trends, policy and strategy consideration*. Paris: UNESCO.
- UNESCO. (2012). *2012 Paris OER Declaration*. Paris: UNESCO. Retrieved from http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/Events/English_Paris_OER_Declaration.pdf.
- Valasek, T. (2001). *Student persistence in web-based courses: Identifying a profile for success*. Raritan Valley Community College, Center for the Advancement of Innovative Teaching and Learning. Branchburg: New Jersey.

- Wladis, C., & Hachey, A. C. (2017). Using course-level factors as predictors of online course outcomes: A multilevel analysis at a U.S. urban community college. *Studies in Higher Education*, 42(1), 184–200. DOI: 10.1080/03075079.2015.1045478.
- Xenos, M., Pierrakeas, C., & Pintelas, P. (2002). A survey on student dropout rates and dropout causes concerning the students in the Course of Informatics of the Hellenic Open University. *Computers & Education*, 39(4), 361 – 377.
- Xu, D., & Jaggars, S. S. (2011). Online and hybrid course enrollment and performance in Washington State Community and Technical Colleges. *Report of Columbia University*, 31, 1–37.
- Yang, D., Baldwin, S., & Snelson, C. (2017). Persistence factors revealed: Students' reflections on completing a fully online program, *Distance Education*, 38(1), 23-36. DOI: 10.1080/01587919.2017.1299561.