A machine learning framework for modelling and optimising patient flow through a hospital setting

Sebastian Gwilym Neale

September 2018

School of Computer Science and Informatics

Cardiff University

Supervisor: Dr Angelika Kimmig

A dissertation submitted in partial fulfilment of the requirements for MSc (in Advanced Computer Science) by taught programme.

Declaration and statement

CANDIDATE'S ID NUMBER	C1573025
CANDIDATE'S SURNAME	Please circle appropriate value Mr / Miss / Ms / Mrs / Rev / Dr / Other please specify ..NEALE.....
CANDIDATE'S FULL FORENAMES	SEBASTIAN GWILYM

DECLARATION

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed .... (candidate) Date ..07/09/18...

STATEMENT 1

This dissertation is being submitted in partial fulfillment of the requirements for the degree of ...MSc...............(insert MA, MSc, MBA, MScD, LLM etc, as appropriate)

Signed .... (candidate) Date ..07/09/18......

STATEMENT 2

This dissertation is the result of my own independent work/investigation, except where otherwise stated.
Other sources are acknowledged by footnotes giving explicit references. A Bibliography is appended.

Signed .... (candidate) Date ..07/09/18.....

STATEMENT 3 – TO BE COMPLETED WHERE THE SECOND COPY OF THE DISSERTATION IS SUBMITTED IN AN APPROVED ELECTRONIC FORMAT

I confirm that the electronic copy is identical to the bound copy of the dissertation

Signed .... (candidate) Date ..07/09/18...

STATEMENT 4

I hereby give consent for my dissertation, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed .... (candidate) Date ..07/09/18...

STATEMENT 5 - BAR ON ACCESS APPROVED

I hereby give consent for my dissertation, if accepted, to be available for photocopying and for inter-library loans **after expiry of a bar on access approved by the Graduate Development Committee.**

Signed .... (candidate) Date ....

Abstract

This project proposes a machine learning framework for modelling and optimising patient flow which can be generalised to most hospital settings at any level of pathway granularity. It is intended that this framework provides a basis for future modelling and optimisation in healthcare settings, providing methods which enable organisations to undertake better capacity and demand planning in a world where healthcare providers need to do more activity to cope with increasing demands, and will need to accomplish this prospect with reduced resource and greater patient and organisational expectations.

The framework proposes four components, datasets, models, simulation and discharge solutions. Datasets are developed relating to admissions, discharges and length of stay using SQL server. Models are implemented using pre-processing techniques and stacked generalisation methods to build accurate models which predict hourly admissions and discharges, and a patients length of stay. Simulation uses the models (and the concept of resources) to simulate patient flow through a specialty. Three discharge solutions are proposed to generate a daily discharge-profile to control and evaluate patient flow. Best practice (80% before 11am), discharge model (based on historical data) and genetic algorithm (which utilises service-level objectives to evolve a good solution).

This project finds best practice discharge solution most effective with increased simulated discharging of 3-12% and a 73-85% reduction to lost patient hours, while genetic algorithm solutions offer increased discharging of 1-12% and a 61-69% reduction to lost patient hours. This project also finds that machine learning models reflect hourly, daily and yearly admission profiles, with most models achieving 75-90% accuracy against test data when predicting a patients length of stay or the number of hourly admissions and discharges. This project finds stacked generalisation methods to be effective when developing demand and activity models, with accurate predictions almost always in the top-two estimators.

Acknowledgments

I would like to thank Brenda Basweti for her endless support and encouragement, without her this work would not have been possible.

I would like to thank my supervisor Dr Angelika Kimmig for her invaluable guidance.

I would like to thank Emma Crawford, Anthony Tracey and Gareth Beynon for their support and context of the problem domain.

Table of Contents

Table of Figures

Table of Tables

Chapter 1. Introduction

The Welsh National Health Service (NHS Wales) restructured during 2009, yet many of its internal organisation are still coming to terms with the impacts of this complex reform where several smaller organisations merged into much larger organisations, with many of its operational processes remaining unchanged. NHS Wales organisations failed to effectively unify their internal services and adapt their service models which has led to poor performance, although the optimisation of all organisational processes are being brought further into the spotlight with each passing year.

This project proposes a machine learning framework for modelling and optimising patient flow through a hospital setting. It is intended that this framework can be applied at multiple levels of granularity (such as an entire pathway, a single ward or specialty, or even a waiting room) where data exists which can be used to derive inbound flow, outbound flow and internal activity. It is intended that this framework provides a basis for future modelling and optimisation in healthcare settings, providing a method which enables organisations to undertake better planning. This chapter introduces the setting background and identifies the specific services evaluated using this framework. An overview of capacity and demand, emergency departments and planning is provided, followed by the project motivation, organisational requirements and project aims.

1.1. Setting

Following a Welsh Government (WG) directed reform in 2009 with the goal of enhancing integrated healthcare in Wales and improving health outcomes, 22 Local Health Boards and 7 NHS Welsh Trusts restructured into 7 Local Health Boards, 3 NHS Trusts and several smaller NHS Wales bodies providing a range of services (Health in Wales, 2009). Recently the Local Health Boards became known as University Local Health Boards except for Powys which is known as a Teaching Local Health Board (Healthcare Inspectorate Wales, 2015) (Welsh Assembly Government, 2009). A key reason for merging smaller local organisations and restructuring into larger organisations – a trend occurring for health organisations across Europe – is primarily a coping mechanism for higher demand and ensuring quality and safe healthcare delivery in the face of a reduced hospital capacity (Sørup, et al., 2013). One such University Local Health Board is Hywel Dda University Local Health Board (HDdULHB) which consists of the three original Local Health Boards - Carmarthenshire, Ceredigion and Pembrokeshire – located in the south west of Wales servicing a population of approximately 380,000. HDdULHB has four acute hospitals providing acute services and many community hospitals, health centres, general practitioners and pharmacists providing a range of services. The four acute hospitals are Withybush General Hospital in Pembrokeshire, Glangwili General Hospital (GGH) and Prince Philip Hospital in Carmarthenshire and Bronglais General Hospital in Ceredigion.

Whilst there is an overlap of specialty services across HDdULHB, each hospital has unique processes which do not generalise to a HDdULHB-wide model, thus this project focuses on GGH which has approximately 18,500 annual admissions and discharges across all specialties and 399 beds. Similarly, each specialty and ward within a hospital has unique processes thus it is decided to limit the scope of this project to three specialties representing three models of patient flow. Originally it was intended to include the three highest volume specialties, however the number of beds could not easily be mapped to some specialties (i.e. numerous sub-specialties such as those occurring in general medicine and general surgery) which would have overcomplicated the models and increased the time demands upon this project. The selected specialties are Urology (URO), Trauma & Orthopaedic (T&O) and Ear Nose and Throat (ENT) which receive a high number of annual admissions and can be mapped to a specific number of beds. Table 1 displays the average number of annual admissions, average LOS and average number of annual discharges for each specialty for the temporal period between April 2013 and March 2018 (five-years).

Table 1: Selected Specialties

Specialty	Average annual admissions	Number of beds	Average length of stay	Average annual discharges
Urology (URO)	1,453	10	3	1,441
Trauma and Orthopaedic (T&O)	2,365	46	8	2,355
Ear Nose and Throat (ENT)	1,827	14	2	1,825

1.2. Capacity and demand

During the past fifteen-years there has been a refocus of staff prioritisation within NHS services with emphasis being placed on making care more efficient, safer and improving patient experience. The refocus is primarily being driven through government policy to improve performance against key metrics such as waiting times for diagnosis, treatment and increased service demand. This is particularly evident for emergency admissions where increasing performance against key metrics is viewed as high-priority (Connolly, et al., 2009).

Healthcare Inspectorate Wales inspections during 2015 identified the need for the HDdULHB to improve patient flow through its hospitals and listed an action to review capacity procedures since they were not deemed fit-for-purpose (Healthcare Inspectorate Wales, 2015).

The lack of capacity and demand planning across Welsh health boards has resulted in the release of several continually evolving frameworks by WG, including the more recent '2018/21 NHS Wales Planning Framework' which outlines the expectations of all NHS Wales organisations through the years 2018-2021. The framework stipulates that all organisations must have a strategy which sets out a long-term vision (including the use of analytics) to inform service, demand and capacity planning and infrastructure decisions. The framework also notes that the public expects the delivery of timely access to care and that each organisation needs to plan how they will meet the existing national targets by using improvement trajectories and setting local expectations through health board wide targets (Welsh Government, 2017).

1.3. Emergency departments

When considering access to emergency services, HDdULHB is prone to many of the challenges faced by healthcare services globally whereby increased demand for access to emergency departments (ED) has resulted in hospital overcrowding. In the UK, this led to the introduction of target-based initiatives to reduce the maximum amount of time between patient arrival and treatment (Sørup, et al., 2013). HDdULHB is bound by the NHS Wales Outcome Framework 2016-17 set by WG which outlines key objectives and targets to improve patient outcomes. The two key measures outlined in the framework which incentivises better patient flow through EDs and – deductively – reduces treatment delays are (Welsh Governemnt, 2016):

- *“The percentage of patients who spend less* than 4-hours in all major and minor emergency care (i.e. A&E) facilities from arrival until admission, transfer *or discharge”*
- *“The number of patients who spend 12*-hours or more in all hospital major and minor care facilities from arrival until admission, *transfer or discharge“*

The former stipulates that 95% of patients (major and minor injuries) should wait less than 4-hours, while the latter is 100% of patients waiting less than 12-hours (both the number and percentage are often used), although all EDs (including those at HDdULHB) in Wales often fail to meet the targets set by WG in the framework (Welsh Government, 2017). Two years of performance data between October 2015 and September 2017 demonstrate that Hywel Dda failed to meet the 4-hour target (Figure 1) and the 12-hour target (Figure 2) on any given month (Welsh Government, 2017).

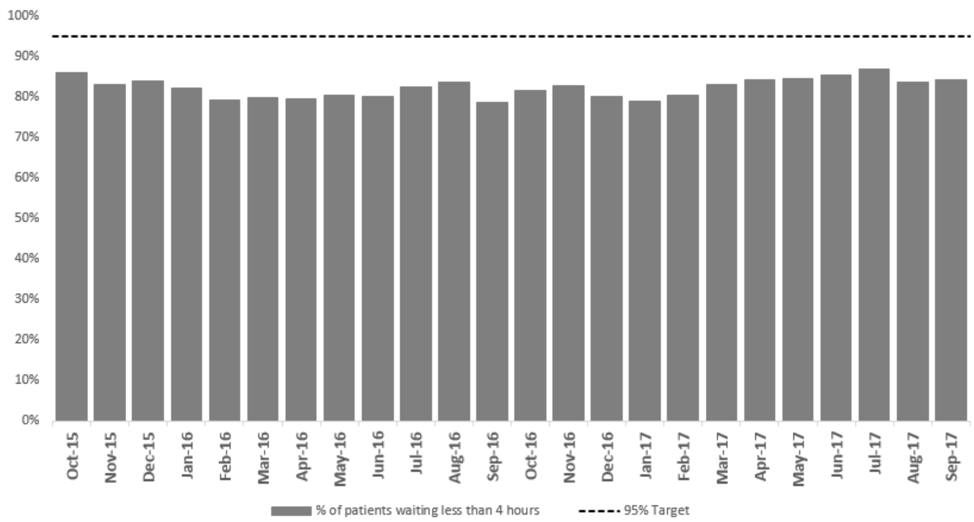

Figure 1: HDdULHB performance against 4-hour target.

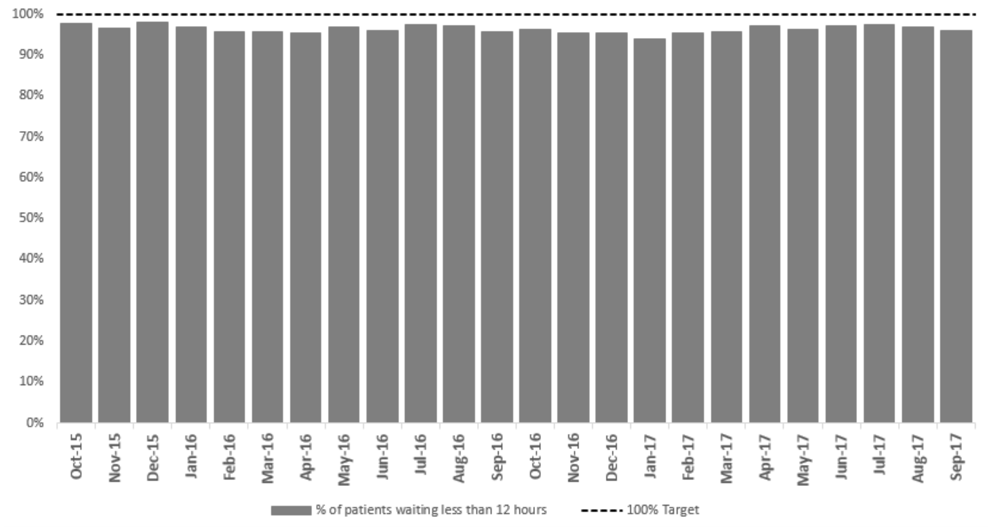

Figure 2: HDdULHB performance against 12-hour target.

1.4. Planning

Providers of healthcare are keen on devising efficient and effective methodologies to solve problems arising in service planning and operational management, this is particularly vital for elements such as staff scheduling, patient flow and discharging which require a balance of conflicting objectives and constrains between service requirements, patient needs, staff needs and the management of services both in-hours and out-of-hours with the former usually being defined as between 8am and 6pm Monday to Friday (Mutingi & Mbohwa, 2013). The desire for enhanced hospital scheduling in what is a complex planning environment is seen as a key factor to help reduce costs (Stolletz & Brunner, 2012), whilst the desire to improve early-day discharging is deemed integral to alleviating the discharge bottleneck to patient flow through a hospital (Khanna, et al., 2016). This issue is caused by late discharges releasing insufficient bed capacity prior to the influx of new patient admissions, causing delays to patient admissions and negatively impacting upon waiting time performance metrics. The desire to improve patient flow has caused many hospital and service planners to desire additional methods to facilitate better demand, capacity and discharge planning (Lin, 2008) (Khanna, et al., 2016).

Another issue for healthcare planning is the need to plan for seasonal variation, particularly winter planning where a combination of issues (i.e. lack of organisation, extended bank holidays, winter weather, increased staff holidays, reduced services) fluctuate both demand and activity, impacting negatively on national performance metrics, staff morale and patient care (Evans, 2017).

1.5. Motivation

The motivation for this project is to improve patient care (through timely access to services and reduced time in hospital), improve staff morale (by lowering hospital crowding and reducing pressure) and increase performance against national targets. The motivation for this project is derived from the need to improve patient flow through a hospital setting by offering additional methods and tools to HDdULHB services which can help facilitate improved capacity and demand processes. By increasing the body of knowledge in the field of healthcare capacity and demand, it should be possible to help HDdULHB to better understand their services from a whole-system perspective.

Effective patient flow targets are not currently set or monitored at a national level in Wales, nor are discharge-profiles used to plan and control daily discharges despite supporting evidence which suggests that early-day discharging can positively affect both patient care and patient satisfaction whilst increasing patient flow. The global healthcare community is starting to appreciate the impact of inter-relationships and inter-dependencies between hospital services, although the relationship with the wider community services and the relationship between hospital admissions, activity and discharges – including its effects at ED – has not been fully explored and more studies are needed to add to the body of knowledge which minimises the effects of a growing issue: an increased hospital demand in an aging population in the face of reduced healthcare resources (Powell, et al., 2012).

1.6. Organisational requirements

HDdULHB believes that effective patient flow through their hospitals is constrained by a failure to meet best practice discharging, prompting some senior managers to ask questions about whether alternative (or more feasible) discharge-profiles could function relatively well compared to those considered best practice, enabling system pressures to be alleviated more quickly and provide a step in the right direction. HDdULHB does not believe that current ED metrics are fit-for-purpose in the context of improving patient care and increasing patient flow, thus HDdULHB is keen to establish new metrics, systems or methods which could focus efforts to improve flow through a hospital.

HDdULHB has provided the following organisational requirements (OR):

OR1. Investigate the use of machine learning to improve patient flow in a hospital setting.

OR2. Identify new metrics which could be modelled through machine learning to improve patient flow in a hospital setting.

1.7. Aims and objectives

The aim of this project is to use machine learning methods and technologies to investigate suitable methods for improving patient flow through a hospital setting while maintaining a balance between the requirements of services and patients.

The following objectives are identified:

1. Identify a key capacity and demand metric from within the literature which will improve patient flow through a hospital setting.
2. Identify technologies which can be used to model patient flow through a hospital.
3. Create a framework which models and optimises patient flow through a hospital setting using technologies identified in objective 2.
4. Evaluate patient flow through a hospital using the metrics identified in objective 1.

At this stage, objectives 1 and 2 are exploratory since HDdULHB has provided a broad set of requirements. These objectives are updated in chapter 3 due to the identification of key metrics and technologies in the literature.

Chapter 2. Review of existing literature

To better understand the problem facing HDdULHB and to establish potential solutions, the concepts of hospital admissions, capacity and discharges is first reviewed, followed by a review of machine learning modelling methods and technologies, and genetic algorithms. Finally, a conclusion of the literature is provided to summarise key findings.

2.1. Hospital admissions

The number of hospital admissions in the UK is increasing year-on-year with the number of inpatients in Wales increasing between 1958 and 2010 (although the rate of increase has slowed since the mid-1990s) (Figure 3), observing a consistent rise in emergency admissions during the same period. The growth of emergency admissions appears to be caused by new attendances since the number of re-admissions has decreased since its peak in the 1960s (Figure 4) (Connolly, et al., 2009) (Statistics for Wales, 2011).

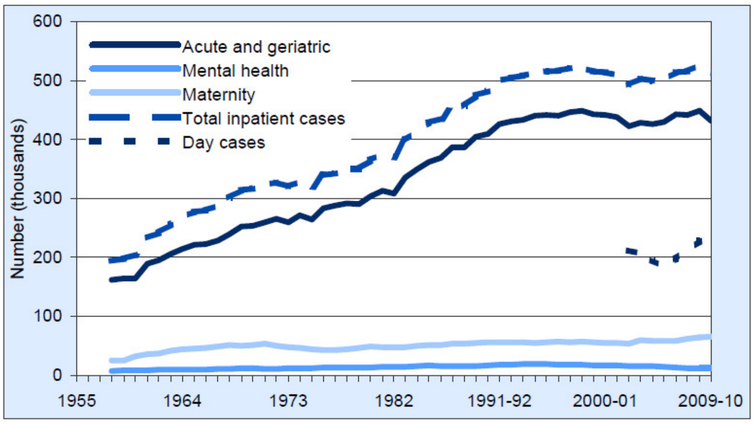

Figure 3: Hospital inpatient attendances (Wales), 1958-2010

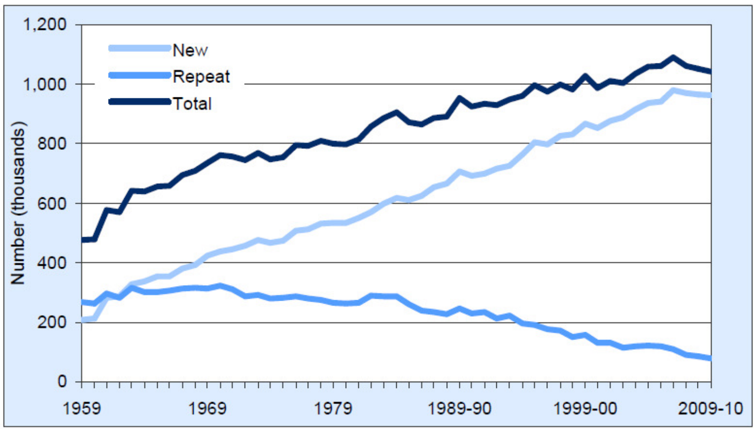

Figure 4: ED new and repeat attendances (Wales), 1959-2010

A national target was introduced by WG to reduce total patient ED waiting time for treatment with the premise that 95% of patients should not spend more than 4-hours, and 100% less than 12-hours before either being admitted into hospital, transferred or discharged. Exceeding the target is known as a breach while some organisations go further and measure 8-hour breaches. (Healthcare Commision, 2007). The rationale for applying this measure is to reduce ED overcrowding and facilitate better patient flow, although it is argued that a single temporal metric does not correlate to high quality of care nor patient safety and can lead to the rise of dysfunctional processes to circumnavigate the targets. An example is the rapid transfer of patients to another ward – sometimes a temporary or holding ward such as a medical assessment unit – often near the time of breach to ensure a patient remains within the target (Sørup, et al., 2013) (Smith, et al., 2014). While the argument against a single temporal metric is valid, a review of ED crowding found evidence to suggest that longer waits are detrimental to the perceived quality of care by patients and that abandonment rates are higher when waiting times are longer. The authors also found that patients are more likely to have a poorer sickness outcome where waiting times are longer (Yeh & Lin, 2007).

To minimise waiting times, it is argued that there should be an increased focus on hospital service associations since correlations are identified between a patients required specialty, the hospital location (both hospital layout, and physical location) and temporal attributes (such as hour of the day, weekdays) impacting upon patient outcomes. Evidence suggests that this issue could be alleviated with the introduction of a scoring system to calculate the likelihood of a patient breach, identifying the patients likely to wait longest. This may reduce the number of breaches and reduce waiting times, although there is also a call for the introduction of adequate flexibility by many clinicians which allows for the exclusion of patients from breach targets where they cannot be avoided, such as the patient being too unstable for admission which suggests that the current metrics and methods may not be fit-for-purpose (Smith, et al., 2014). Whilst it is reasonable to believe that there would be more breaches on a busy day than a quieter day, analysis on 4-hour breaches at Salford Royal Foundation Trust did not find a significant difference between the number of breaches on a non-busy, nor a busy day – with busy being defined as above the 75th centile for daily patients in the previous six months which suggests that the problem of ED crowding is not caused by service demand in isolation (Smith, et al., 2014).

A key factor at play with extended ED waits is increased hospital occupancy and reduced bed capacity which results in lower bed availability. This leads to a crowding of patients waiting for admission from ED and results in a worsened patient experience, increased treatment delays, a lower quality of care and a rising number of potential errors. There is also evidence that ED staff are distracted by the flow of new patients into ED since the nature of the department is often to deal with urgent cases, a problem which is exacerbated as ED occupancy is increased (Liu, et al., 2009).

A key bottleneck has been identified at ED; the rate at which patients admitted to ED leave the department for other inpatient beds within the hospital. It is believed that by improving the rate of flow to other hospital wards, there should be reduced waiting times and breaches at ED through freeing up ED beds, leading to improved patient experience and enhanced patient care (Powell, et al., 2012).

2.2. Hospital capacity

The number of hospital beds in Wales has decreased 25% between the year 2000 and 2015 with a rate of 5 beds per 1,000 people in 2000, and 3.5 per 1,000 people in 2015. Bed occupancy has risen in the same period to record rates of 86.9% (Figure 5), although average length of stay (ALOS) has decreased from 8 days to below 7 days (British Medical Association, 2016).

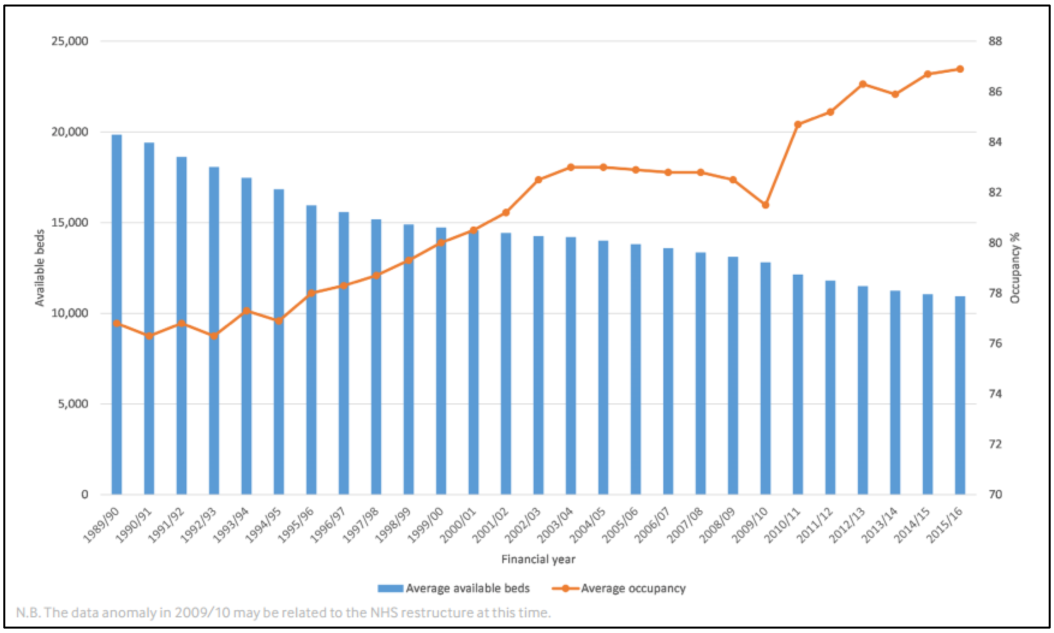

Figure 5: Number of beds vs. bed occupancy (Wales) 1989-2016

When reviewing literature for an optimisation project, Yeh & Lin (2007) found that hospital bed shortages are a source of overcrowding at ED. The understanding is

that activities outside the control of ED (such as inpatient bed management, elective admissions and discharge activity) impacts upon the flow out of ED into other appropriate locations of care, contributing to ED crowding. It is believed that hospitals with high occupancy rates (which results in constrained and inflexible capacity) need to optimise both patient flow and the utilisation of resources, however this is often undertaken operationally by management teams who do not always provide enough consideration to the holistic approaches of care which can result in disjointed services being offered to the patient. This does not mean that management teams do not understand the requirements, it is more a symptom of the pressures faced by the healthcare systems (Khanna, et al., 2016). The consequences of ineffective hospital capacity and flow management includes patients being admitted to clinically inappropriate wards, a reduction of patient satisfaction, reduced staff morale and adverse patient care (such as increased infection rates or patient distress, particularly amongst older patients). Uncontrolled flow has also raised concerns amongst healthcare professionals where there is pressure to free up beds, resulting in discharging patients before it is safe to do so (British Medical Association, 2016).
One method to reduce capacity constraints during times where ED admissions are under strain due to crowding is to reschedule elective patients (which negatively impacts elective patient care and referral targets) or to admit patients to temporary locations (including hallways, temporary assessment units or in extreme cases, remaining on an ambulance longer than necessary), all of which does not solve the capacity management issues since the demand still exists and must be met at a later stage (Powell, et al., 2012). A solution to this complex problem of hospital capacity management has been suggested for several decades with the belief that use of simulation could enhance the management and structuring of many departments within a hospital setting, including ED (Yeh & Lin, 2007).
Another suggested solution to capacity and demand management is to balance the number of patient discharges with the number of new admissions to ensure that there is enough hospital capacity to meet incoming demand. This should facilitate better patient flow through a hospital and reduce the amount of time patients wait at ED since there is a known correlation between the increasing hospital occupancy rate and the increasing amount of time patients wait at ED, although this is dependent upon there being specific specialty or community capacity available to meet the patients' needs (Powell, et al., 2012). To achieve this goal, the hospital management

teams would require some form of capacity planning tool which utilises demand analysis to decide what capacity requirements and activity levels are needed to meet demand, and a method for establishing the number of discharges is needed to ensure there is sufficient capacity in the system (Subramanian, 2013).

Another key issue with the management of hospital capacity is staff scheduling with Puente et al. (2009) stating that an effective staff schedule should comply with organisational policies and objectives, but also balancing resources, the delivery of quality healthcare and staff preferences which can often impact upon staff morale (Puente, et al., 2009). Research by Yeh & Lin (2007) using simulation models found that by optimising nurse schedules, it is possible to reduce patient queuing times by 43% without increasing the number of nurses in the system (Yeh & Lin, 2007).

2.3. Hospital discharges

Hospital discharge planning is a fundamental yet often overlooked aspect of healthcare planning in hospital settings despite a consensus that discharge planning should begin at the patients admission stage (Maramba, et al., 2004).

While hospital admissions are increasing, LOS is decreasing (Figure 6) due to legislation and policy incentivising a focus upon early discharge and reducing delayed transfers of care (DTOC) without fully considering the complexities and obstacles associated with discharge planning both in hospital and community settings (Connolly, et al., 2009) (Statistics for Wales, 2011). By targeting one service area, unintended consequences are occurring in others, for example by reducing LOS through hastening a patients discharge, there is a reduction of time available for nurses to meet the patients post-discharge needs. During a patient experience survey in 2007, hospital discharging (including the corresponding planning and coordination of discharges and safe discharges) is fifth in the top-ten complaints referred to the healthcare commission which violated a key aim of healthcare – providing patient safety. A lower ALOS can also result in the negative effect of increasing hospital re-admissions (Healthcare Commission, 2008) (Maramba, et al., 2004).

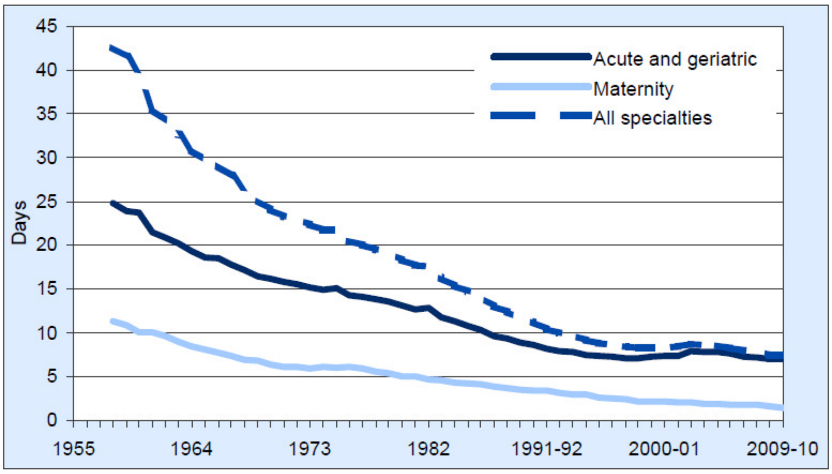

Figure 6: Average length of stay (Wales) 1958-2010

When considering constraints to the discharge process, research conducted by Connolly et al. (2009) identified two key themes, ‘conflicting pressures’ and ‘casualties arising from conflicting pressures’. The authors found that conflicting pressures associated with discharging patients, particularly when treating patients with more complicated needs is apparent, as is the pressure relating to the term medically stable which encouraged a rapid trajectory towards discharge despite other factors existing which made the push to discharge either unsafe or difficult to achieve. This problem is exacerbated when the availability of outside services is limited and further aggravated by processes which are deemed to be inflexible to the patients’ needs. While a patients family are typically called upon to assist with the discharge process, it was found that they sometimes do not cooperate for various reasons including financial or schedule issues which further disrupts the process (Connolly, et al., 2009). Connolly et al. (2009) found that processes used to alleviate many organisational barriers seem to break down due to the complexity of health systems and its relationship with external organisations. i.e. while some parts of the organisations tried to address issues with weekend discharges, other elements of the organisation stalled progress (Connolly, et al., 2009). A patient experience survey in 2007 found that 38% of patients experience a delay on the day they are discharged, while a survey of independent sector treatment centres found that only 14% of patients experienced delays on the day of discharge; however such centres usually served patients requiring elective procedures which are usually less complex than emergency treatment (Healthcare Commision, 2007). There is also evidence to suggest that factors such as postcode can influence the level of aftercare received, particularly where some local authorities are better coordinated, accessible and offer strong communication links with hospitals (Connolly, et al., 2009).

Location coordination is another issue with patients not being in the correct location at the required time of discharge (Allen, 2001), although there is evidence that deployment of a discharge coordinator could alleviate this issue and enhance the discharge process by following patients throughout their hospital stay and facilitating communication between teams, departments or external organisations. A discharge coordinator would facilitate the planning and execution of patient transport processes (i.e. contacting the family, arranging local authority collection), coordinates medication and patient notes in preparation for discharge, and contact the patient for follow-up information (Connolly, et al., 2009) (Maramba, et al., 2004). The preparation of timely medication could improve the process as 61% of patients who took part in an inpatient survey stated that they experienced a discharge delay due to delayed medication. (Healthcare Commision, 2007).

The reduction of resources within community settings has also strained the discharge process with a rise of DTOC into the community over the past three years, although as of March 2017, 55% of discharge delays are credited to problems originating within the NHS (Care Quality Commission, 2017) which is in contrast to delays relating to older patients in a 2003 report by the National Audit Office which found that the majority of DTOC are caused by external issues (Bourn, et al., 2003). In Wales, the number of patients experiencing DTOC decreased between June 2004 and December 2010 (Figure 7) (Statistics for Wales, 2011).

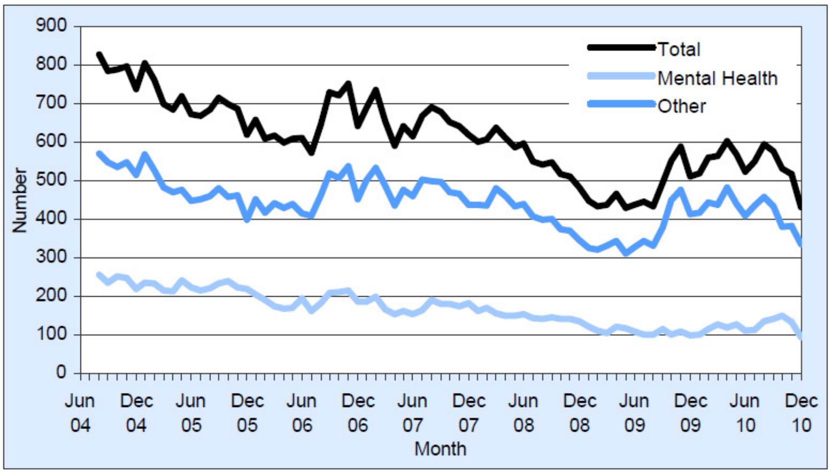

Figure 7: Delayed transfers of care, Aug-04 to Dec-10

The impact of ineffective discharging upon hospital flow has increased the desire for new and effective discharge solutions. One of which is a discharge target where 80% of patients discharged before 11am is deemed to be popular by academics and management teams, although less so by clinicians. Simulations conducted by Khanna, et al. (2016) to establish optimal discharge target found that 80% before 11am could result in a 25% reduction in bed wait times and allowed an additional 1% of patients to flow through the hospital (Khanna, et al., 2016). The study makes a stronger case for early-day discharges to improve patient flow rather than staged discharges, while noting that different strategies and distributions may suit some hospital services more than others. Some clinicians have called for a discharge target which can be spread out across the day to match incoming demand, deeming early-day discharges not operationally realistic. The authors state that based upon their research and review of the subject area, there is no evidence to facilitate the creation of an optimal staged discharge target for hospital services which provides potential for further study using alternative techniques (Khanna, et al., 2016). The authors also found that a later-day staged target can be almost as effective as early-day discharging but remained less effective which reinforced their belief that most discharges should occur within the first half of the day (Khanna, et al., 2016). Where there has been a successful deployment of uniform daily discharges models through simulation, it was found that discharging at an interval of 1-hour earlier (shifting the distribution curve forward by 1-hour) results in reduced patient waiting times at ED and improved bed availability, while discharging 4-hours earlier reduced all waiting time breaches. (Powell, et al., 2012).

2.4. Machine learning modelling

Data mining (DM) is defined as the process of extracting and discovering meaningful patterns in data which can be used to offer some form of advantage, with the main tasks including classification (mapping data inputs to a class), clustering (identifying a finite set of clusters) and association (finding relationships between attributes). DM makes use of several fields including machine learning (ML), statistics and databases. ML is a term for a collection of tools and techniques which allows users to make use of what is becoming an overwhelming amount of data by learning from it (Witten, et al., 2017).

Stacked Generalisation (SG) is an ensemble method used for classification where several ML algorithms are combined to improve the performance of predictions. SG consists of individual level-1 generalisers (often several classifiers or regressors, each generating unique models) and a level-2 generaliser generating a model of model predictions (Wolpert, 1992) (Tugay & Oguducu, 2017). This technique has been used to forecast hourly sales demand for an online e-commerce company, using stage-1 regressors (decision tree, random forest, gradient boosting and linear) which are fed forward into a level-2 meta-regressor. Results demonstrate that the model predicted demand at least as well as individual models, often better with only 20% of the data was required for the same results which led the researchers to believe that with more data, the model could be more accurate. The authors state that whilst there was a notable performance improvement, there is no statistical significance between the results (Tugay & Oguducu, 2017) however improved accuracy alone is not the only reason to utilise SG since research suggests that SG models will benefit from improved generalisation accuracy as opposed to learning accuracy (Wolpert, 1992). The SG model proposed by Tugay & Oguducu (2017) could also benefit from replacing the least accurate algorithm with alternative algorithms which are known to perform well for forecasting demand in other fields. A multi-layer perceptron (MLP) is a feed-forward artificial neural network (ANN) which consists of an input layer, one or more hidden layers and an output layer, using backpropagation to update weights in earlier layers. MLP has been used to predict hourly demand for energy consumption and load forecasting with accuracy surpassing traditional ML and time series regression models, although one issue cited using MLP for forecasting is overfitting (Rodrigues , et al., 2014) (Ryu, et al., 2016) (Szoplik, 2015) which could be reduced by including it as part of a more generalised SG model.

A key issue faced when developing ML models is imbalanced datasets, with imbalances of up to 100:1 common (i.e. fraud detection since fraudulent activity is less common that non-fraudulent activity) although some applications and fields can find imbalances of up to 100,000:1 (Chawla, et al., 2002). With respect to hourly hospital demand at lower levels of pathway granularity, the number of 0 and 1 patient attendances will be higher than the number of 2 or more, potentially causing imbalanced data and underfitting models. A solution to this issue is the synthetic minority oversampling technique (SMOTE), where new samples of minority classes are synthesised using clustering techniques as opposed to oversampling through replacement, or undersampling through removal from the majority class. SMOTE has been proven to enhance classifier accuracy (particularly where there is low-dimensional data) based upon experiments which compared alternative solutions, one such experiment utilised nine unique datasets. A key issue with SMOTE is that overfitting models can occur since there is a bias towards the minority class, although this could be mitigated or reduced by utilising SG methods (Chawla, et al., 2002) (Blagus & Lusa, 2013).

A final note for demand modelling is that the application of hospital forecasting through simulation is an area of interest to service managers, although few studies exists which can be generalised to other hospitals (Jones, et al., 2009). A systematic review of models for forecasting the number of ED visits concluded that a calendar model for prediction is the easiest to understand, particularly for professionals with no statistical background (Wargon, et al., 2009). Demand prediction in healthcare using temporal attributes of hour, day of week, month of year and to a lesser extent, public holiday are deemed to be effective (Xu, et al., 2013), with other studies expanding upon these temporal attributes, adding elements such as season, historic weather events and rolling average demand to predict variability (Hertzum, 2016).

2.5. Genetic algorithms

Genetic algorithms (GA) can be described as a simplified form of evolutionary algorithms inspired by the term “survival of the fittest” to imitate the concept of “natural selection” through a random yet directed search through a solution space. GA began in the 1960s as a solution to optimisation problems by using adaptive search algorithms which are often subdivided into the two classes, deterministic or probabilistic. GA are essentially a probabilistic approach to solving an optimisation problem (Subramanian, 2013). When creating GAs, individuals (problem solutions known as chromosomes) are encoded as strings just as chromosomes in the field of genetics. For each optimisation problem, an individual represents a specific point in the search space of all potential solutions. Since the population represents a set of possible solutions, the fittest (evaluated as most suitable through a fitness function) are more likely to be selected and generated into new solutions using the genetic operators of ‘selection’, ‘crossover’, ‘mutation’ and ‘replacement’ which replaces the previous generation with a new generation consisting of better features. The process of selection and reproduction continues until some stop criteria is met, producing increasingly optimal solutions (Puente, et al., 2009). GA are particularly useful for problems relating to scheduling, trajectory planning, operation control, signalling processing, designing neural networks and travelling salesman problems. GA are particularly effective where there is a need to maximise or minimise an attribute which is subject to several constraints, and where the problem solutions are deemed NP-Hard (Subramanian, 2013) (Mutingi & Mbohwa, 2013).

While GA are not used to optimise flow within healthcare settings, they are often used to optimise supply chain activities by minimising total cost through balancing constraints against production levels, stock levels, labour costs (including overtime costs) and demand. Research by Subramanian (2013) found that using GA techniques provides good quality solutions to the problem in that the outlined conditions are met, although noted that the solutions may not be the best since optimal solutions cannot be created using only evaluations of a polynomial function, particularly where there is a larger range of possible solutions for effective non-directed search. The researcher also described some of the limitations which include the difficulty of representing accurate constraints relating to real-life planning and time planning problems in a search algorithm, yet despite such limitations, Puente et al. (2009) argues that use of GA is increasing due to the proven ability to evolve near-optimal solutions to optimisation problems across a variety of fields which are non-linear in nature (Subramanian, 2013) (Puente, et al., 2009).

The use of multi-objective GA offers the potential to generate discharge-profiles based on hospital demand, activity and conflicting constraints such as patient considerations, staff considerations and prioritisation. Whilst this specific method has not been used for healthcare pathway optimisation, the approach has been used to generate staff rotas, balancing the required staffing levels to manage hospital demand with individual (often conflicting) staff preferences whilst offering perceived fairness and managing uncertainties (Mutingi & Mbohwa, 2013) (Yeh & Lin, 2007). This technique was used to satisfy staff preferences by considering their needs and could be used in a similar way when generating discharge-profiles which balance the constraints by maximising discharges, maximising patient needs, minimise hospital occupancy or maximising patient flow.

Puente et al. (2009) encountered limitations when deploying constraints to optimise doctor rotation schedules since it was difficult to balance the essential service and organisational requirements with the wishes of the staff where there is equal weighting. Thus Puente et al. (2009) introduced the organisational requirements as hard-constraints (typically characterised by coverage, such as the minimum and maximum staff required for a shift) which all need to be fulfilled, and soft-constraints (typically characterised by personal staff schedule needs) which should be met as much as possible. The aim was to schedule resources which met all the hard-constraints while maintaining a high quality of soft-constraints and was achieved by providing a solution scoring value for each of the attained soft-constraints, with the sum of the values acting as the fitness function. This ensured that the most important constraints were rarely unsatisfied (Yeh & Lin, 2007). The development of effective constraints requires a translation of constraints into a metric (such as number of weekly shifts, or total number of discharges) which can be used to assess feasibility of the solution and where a constraint is violated, apply penalty functions (such as quadratic distance function) to evaluate the fitness of an individual (DEAP, 2018).

2.6. Literature conclusion

The number of new hospital admissions is increasing, the number of repeat admissions is decreasing, and the growth of inpatients has slowed, although the rate continues to rise steadily. There is no evidence that current targets in isolation improve patient flow through a hospital which appears to be constrained by the lack of capacity outside of ED, both within the hospital and in the community. Clinicians believe that patients should be provided with flexibility to remain in ED without a strict time limit, and that a scoring system could be introduced to help identify patients at risk of breach at an earlier stage. There is also a desire to reduce ED crowding (which is a cause of distraction for staff and leads to a low quality of care) by transferring or discharging patients to more appropriate care locations.

In contrast to increasing admissions, bed numbers in Wales is decreasing, bed occupancy is rising and there is a feeling of increased pressure despite a reduction to ALOS. A combination of unmanaged patient flow and restricted capacity is constraining healthcare services, with some mechanism needed to manage the flow of patient out of hospital. It is evident from the literature that the current methods for managing patient flow through a hospital setting using ED waiting targets in isolation is not fit-for-purpose, particularly when combined with capacity management techniques which facilitates or incentivises inappropriate bed management, the rescheduling of elective patients, transferring patients to temporary locations and unsafe discharging. Many of the existing processes designed to encourage rapid discharge are often inflexible to individual needs, cause patient dissatisfaction, increase readmissions and are neither staff nor patient centric. This is an important point to consider and should be a key factor when designing new discharge processes or metrics, as is the availability of external capacity and external factors such as family cooperation, external capacity and pre-discharge requirements such as medication which sometimes impact negatively upon the discharge process, although it is worth considering that most discharge delays are attributed to problems originating within the NHS. A key solution to the discharge management problem is the establishment of effective discharge targets with 80% of patients before 11am being identified as effective, however there is a call for discharge targets to be distributed throughout the day based on demand and specialty – with no evidence to suggest what the characteristics of an optimal staged discharge-profile for hospital services should be. The insight that 80% of patients discharged before 11am is effective, along with models and simulations which found uniform daily discharge shifts of both 1-hour and 4-hours to be effective offers some direction for further work, as does the deployment of discharge coordinators. Generally, there is a need to optimise patient flow and several solutions are proposed in the literature including the use of simulation to understand and optimise services, the use of more effective capacity management techniques by balancing discharges with demand (preferably with hospital specialty being considered), and the use of improved staff scheduling systems. DM and ML for demand modelling is proven in several fields although appears to be underutilised for demand and activity modelling in healthcare settings. There are several tools and techniques available which can be applied to predict demand and activity of non-linear classes (classes which are not linearly separable), however a key concern with some methods is model overfitting, or underfitting caused by data imbalance. While hospital demand is typically viewed as predictable and static, it is unlikely that the exact number of patients will arrive for a given hour in future as the same temporal properties relating to a historical hour. Demand and activity modelling in healthcare can benefit from balanced generalised predictions to forecast the general trend throughout a year which could be achieved by utilising existing methods such as SG, but also adding additional methods such as MLP and SMOTE.

GA are a probabilistic approach to solving optimisation problems based on the concept of natural selection and can be used to derive a good (although not always the most optimal) solution where there is a need to maximise or minimise an attribute (such as cost or time). The use of GA to generate staffing schedules which balance multiple objectives, or to minimise patient waiting times and staff costs in ED have been achieved using hard-constraints, soft-constraints, quality objectives and combining GA with simulation tools. An important aspect for the success of a GA is the conversion of constraints into a metric which can be minimised or maximised, and where necessary, assigned a penalty. GA are also used as a solution to the demand, capacity and activity problems of supply chain management which could also be applied to demand and capacity patient flow.

From the literature, there is enough evidence that a machine learning framework could provide a system which helps facilitate the flow of patients through a hospital setting. There are several known factors which can delay the discharging process as well as many patient and service level requirements to establish a balanced discharge-profile. By creating a solution which incorporates hourly discharge slots throughout the day and converting some of the conflicting demands into numerical constraints which can be minimised, it is possible to prioritise patients to suitable slots and maximise flow. ML classification models using SG models have been used to predict demand and could be used to form the prediction aspects (in relation to admissions, LOS and discharges) of a simulation tool to generate and compare optimised discharge-profiles. GA have been used to solve demand and capacity issues in supply chain management and improve staffing levels and patient flow through ED. It is possible that GA could be used to create a method which generates an optimised discharge-profile based on conflicting objectives, however such a tool which offers the minimum required number of discharges per hour will need to be utilised in conjunction with a deployment method for effective use. One possible approach to achieve this is the use of one or more discharge coordinators who manage the patients discharge process as safely as possible.

Chapter 3. Problem statement and approach

This chapter provides the problem statement, framework components, project approach and evaluation methods. The problem statement includes updated operational requirements and research objectives following review of existing literature which is discussed and agreed with HDdULHB. The framework components section provides the key components which form the framework, and the project approach outlines the methods considered to generate datasets and models including granularity decisions before outlining the simulation and discharge solution methods used. Finally, the evaluation methods section provides the key questions (and corresponding tests) which are be applied as part of this projects evaluation.

3.1. Problem statement

Chapter 1 and 2 identifies key components of the problem solved by this project and can be partly summarised with the statement: with increasing demand and reduced capacity HDdULHB is expected to do more with less. There are many new and existing pressures exerted upon the internal systems of healthcare providers which is causing fragmented care and insufficient consideration to the needs of its staff, patients, temporal attributes and the unique constraints faced by each specialty. ED crowding is a concern and HDdULHB is failing to meet national targets with a view that ED crowding is often caused by a lack of capacity elsewhere in the hospital system, or externally in the community. It has been established that the current targets are not suitable in isolation and there is a desire to optimise patient flow with performance measured against new metrics whilst balancing the requirements of the service (including clinicians, operational, management etc.) with the needs of patients. There is a desire to deploy better demand and capacity tools and techniques which facilitate better patient flow, but also remove some of the pressures associated with providing rapid discharges in place of safe discharging without ignoring the complexity of discharge obstacles. Efficient discharging is seen as a key to unlocking some hospital flow bottlenecks, with failure to confront the problem leading to patients being treated at inappropriate locations relative to their healthcare needs.

The use of simulation is viewed as a potential solution to facilitate better management decisions and operational structure, with calls for capacity and demand planning tools which facilitate better planning including systems such as those which establish (and balances) the number required hourly discharges (i.e. balancing the number of discharges to the number of admissions) as opposed to an inflexible discharge target as a percentage. With these elements in mind, it is worth being clear that discharging 80% of patients before 11am is currently considered the optimal solution, thus any proposed discharging solutions will need to be evaluated for performance against this best practice discharge-profile.

3.1.1. Revised OR requirements

These complex problems cannot be all solved by this project, but improvements can be made through better understanding if fit-for-purpose concepts are applied. This project will provide a machine learning framework with corresponding foundation systems to help HDdULHB understand its services in greater detail and understand the impact of discharging upon patient flow based on historical data, thus provide a mechanism to manage and optimise future patient flow.

Based upon potential solutions arising in the literature, HDdULHB has agreed to revise the OR requirements to the following:

- OR3. Develop a machine learning capacity and demand modelling framework which facilitates the optimisation of patient flow through an element of a patient pathway, subject to service constraints.
- OR4. Develop a system which simulates patient flow through an element of a patient pathway, utilising data models derived from data within the HDdULHB data warehouse.
- OR5. Develop a system which provides the number of patients to be discharged by hour of the day based upon the number of patients awaiting discharge.
- OR6. Evaluate the effectiveness of alternative discharge-profiles when compared to:
    - 6.1. Best practice discharge-profiles.
    - 6.2. Historical discharge-profiles derived from data within the data warehouse.

OR3, OR4 and OR5 will satisfy OR1, while OR6 will satisfy OR2 in chapter 1.6.

3.1.2. Revised objectives

Considering the objectives in chapter 1.7, objective 1 is achieved since a daily discharge-profile is identified as the capacity and demand metric to improve patient flow. Objective 2 is also achieved since ML tools and techniques, including GA are identified to model patient flow through a hospital setting. The following sub-objectives are now included to form objectives 3 and 4, including the corresponding mapping to OR3, OR4, OR5 and OR6:

3. Develop a framework which provides the structure of a model representing patient flow through a hospital setting (OR3).
    - 3.1. Develop a system for extracting data as features aggregated to an hourly level of granularity for three specialties from the HDdULHB data warehouse (OR4).
        - 3.1.1. Establish the data extraction methodology.
        - 3.1.2. Establish which three specialties are included in the experiment.
        - 3.1.3. Establish which features are used.
        - 3.1.4. Develop a patient admission dataset for each specialty
        - 3.1.5. Develop a patient LOS dataset for each specialty
        - 3.1.6. Develop a patient discharge dataset for each specialty
    - 3.2. Develop a system which takes a dataset outlined in objective 3.1 and generates an optimised model which for a set of features (objective 3.1.3) predicts a target (OR4).
        - 3.2.1. Develop a model which predicts the number of patients expected to be admitted for a specific hour of the day.
        - 3.2.2. Develop a model which predicts the LOS of each admitted patient.
        - 3.2.3. Develop a model which predicts the number of patients expected to be discharged for a specific hour of the day.
        - 3.2.4. Optimise each model developed as part of objectives 3.2.1-3.2.3 to improve the accuracy of the model predictions.
    - 3.3. Develop an application which converts the number of patients waiting to be discharged into a discharge-profile subject to one or more operational constraints (OR5).
        - 3.3.1. Establish suitable operational constraints.

3.3.2. Develop a suitable method for generating a discharge-profile subject to the constraints identified as part of objectives 3.3.1.

3.4. Develop an application which simulates patient flow through a specialty (OR4)

3.4.1. For each hour of the day, predict the number of new admissions using the models created as part of objective 3.2.1.

3.4.2. For each new patient, predict the patients LOS using the model created as part of objective 3.2.2.

3.4.3. For each hour of the day, obtain the number of discharge slots available using one of the following solutions:

3.4.3.1. The discharge model created as part of objective 3.2.3.

3.4.3.2. The best practice discharge-profile identified in the literature.

3.4.3.3. The application created as part of objective 3.3.

4. Evaluate the discharge solutions in relation to the following metrics (OR6):

4.1. The number of lost patient hours (the number of hours between the earliest possible hour of discharge, and the actual hour of discharge) at the end of the simulation.

4.2. The total number of patients discharged (flow) at the end of the simulation.

3.2. Framework components

Based upon the OR requirements, objectives and analysis of literature, the scope of this project is to create a framework which generates datasets, converts datasets into prediction models, then uses the models as part of a simulation where several discharge solutions generate and evaluate discharge-profiles. Figure 8 outlines a high-level view of the interaction between components, with the following components forming a framework:

1. Datasets - A method for extracting datasets relating to admission, LOS and discharge.
2. Models - A method which converts a dataset into an optimised model which for a set of features (data attributes), predicts a target (hourly admissions, los, hourly discharges).
3. Simulation - A system which uses the models to predict the number of hourly admissions, and LOS for each new admission. Then uses a discharge-profile to establish discharge slots.
4. Discharge solutions – Provides a discharge-profile:

- Best practice profile: A discharge-profile based upon best practice discharge solution (what the service should be doing).
- Discharge model profile: (what the service has been doing).
- Genetic algorithm profile: A discharge-profile which allows the utilisation of service constraints as inputs (what the service could be doing).

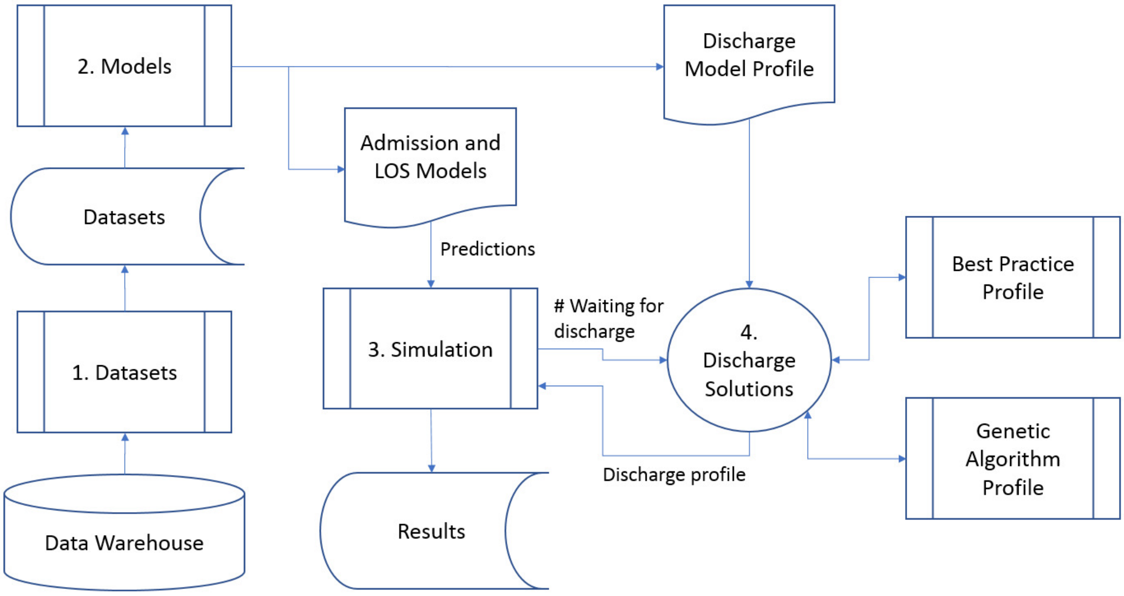

Figure 8: High-level interaction between framework components

3.3. Project approach

This section outlines the approach used in relation to research methods, temporal considerations, demand and capacity models, simulation and proposed discharge solutions. This chapter also provides the evaluation methods used to evaluate the project.

3.3.1. Research methods

The classification of this research is quantitative research since real quantifiable data is being used, and evaluations are performed using quantifiable performance metrics.

The nature of research is analytical since historical data is obtained, based upon the facts of real patient spells in a hospital setting, although the purpose of this project is to provide applied research, bringing together known methods in the field of machine learning to provide a framework which can be adopted and applied in a healthcare setting. The methods used are discussed in the next few chapters.

3.3.2. Datasets

To develop component 1, three themed datasets are extracted from the admitted spells dataset (ASDS) and converted into a dataset of features. The themes are admissions, LOS and discharges, with each dataset replicated for each selected specialty. For the remainder of this report, these datasets will be referred to as admission-dataset, los-dataset and discharge-dataset. Whilst specialty datasets can be grouped using this framework, specialty processes will vary meaning that any generalised model may not be representative to the unique specialty. An initial exclusion criteria is applied at the point of extraction by excluding patient spells where there is a known conflict with the intended purpose of the project. i.e., one exclusion is outpatients since the project is only concerned with patients who are admitted into a specialty and stay overnight, thus the methodology only includes inpatients with a LOS of >=1. The rationale for this is that unknown same day discharges are difficult to stage throughout the day thus this project only considers patients who are ready for next day discharge where there is suitable time for planning. The complete dataset methodology is available in Table 2 which applies HDdULHB standard reporting methods. The complete data conversion is available in Table 3 which utilises some HDdULHB standard reporting conversions (i.e. LOS categorisation), and attribute categorisation relevant to the project requirements (such as sufficient class representation for cross-validation).

The ASDS contains data relating to each individual patient hospital spells (which may include multiple episodes grouped as one transaction) with attributes such as a patients admission date, patient features (i.e. age and sex), specialty, LOS and discharge date. The data attributes within ASDS are suitable to derive hospital demand (admissions-dataset), hospital activity (los-dataset), establishing discharge patterns (discharges-dataset) and generating a basic patient profile. The dataset includes 95 data attributes and 1.4 million data objects dating back to 1st April 2007, however only 8 attributes are used to derive all features required to form admission-dataset and discharge-dataset, and 10 attributes to form los-dataset.

Table 2: Dataset inclusion and exclusion criteria

Criteria	Type	Rationale
Patients who have passed away	Exclude	The discharge process for deceased patients is different from the standard discharge process.
Inpatients	Include	The experiment is only concerned with patients who can be discharged the next day. Inpatient stay overnight, but outpatients do not typically stay overnight and are not usually formally admitted as a patient.
LOS >= 1	Inclusion	This eliminates the low number of inpatients who do not stay overnight.
Same admission and discharge specialties	Inclusion	Some patient pathways may require that a patient is transferred to another specialty during a spell, but transfers are excluded from the scope of this project to limit complexity.
Patient spells within the inclusion timeframe	Inclusion	Training and test dataset include data between 1st April 2013 and 31st March 2018. For the admission and LOS datasets, the inclusion will be based upon the admission datetime. For the discharge dataset, the inclusion will be based upon the discharge datetime.
Glangwili General Hospital	Inclusion	Whilst specialties operate using varying processes, processes for same specialties may vary across hospitals. Thus, to reduce the complexity of the models, this experiment will only consider specialties within Glangwili General Hospital.
Specific specialties	Inclusion	Urology, Trauma & Orthopaedics and ENT are to be included

Table 3: Dataset conversions

Conversion	Rationale
Partially-grouped admission categorisation	Each admission class may have insufficient class representation to perform effective cross-validation. Attributes with classes fewer in number than 5 will be converted to the class below. i.e. values of 6 (if the count is < 5) will become values of 5. This ensures a minimum of 5 classes per attribute.
Partially-grouped age categorisation	Ages above 95 will be converted to 95. This is due to the number of patients aged over 95 being too low for effective cross validation. This ensures a minimum of 5 classes per attribute.
Fully-grouped age categorisation	In addition to the partially-grouped age category, an additional fully-grouped age category will be generated based upon age groupings used by HDUHB for analysis reports. These groups are: • 0-17 • 18-64 • 65-74 • 75-84 • >=85
Fully-grouped LOS categorisation	LOS attributes may have insufficient class numbers to perform effective cross-validation, and the gap between classes may be too large for equal ranges due to outliers. The following LOS conversion will be performed using categories which are used in some HDdULHB reports: 1 to 3 days becomes 1 4 to 6 days becomes 2 7 to 9 days becomes 3 10 to 14 days becomes 4 15 to 19 days becomes 5 20 to 24 days becomes 6 25 to 29 days becomes 7 30 to 39 days becomes 8 40 to 49 days becomes 9 50 to 69 days becomes 10 70 to 89 days becomes 11 90 to 109 days becomes 12 110+ days becomes 13
Partially-grouped discharge categorisation	Each discharge class may have insufficient representation to perform effective cross-validation. Attributes with classes fewer in number than 5 will be converted to the class below. i.e. values of 6 (if the count is < 5) will become values of 5. This ensures a minimum of 5 classes per attribute.

3.3.3. Temporal

The temporal granularity is 1-hour, beginning at 0 and ending at 23. Core hours are between 8:00 and 19:59, represented as 8 to 19 while non-core hours are 0 to 7, and 20 to 23. There are several accepted definitions of core hours including 9:00-16:59 and 8:00-17:59, however 08:00-19:59 was selected to provide 50% core and 50% non-core hours. There is no differentiation between weekdays and weekends.

There are two independent temporal evaluation methods. The core hours method (Figure 9) which considers discharge-profiles during core hours only, and the all-hours method (Figure 10) which considers a discharge solution discharge-profile during core hours, and a discharge model discharge-profile (historical) during non-core hours. The purpose for differentiating between methods is that a controlled discharge-profile only discharges within core hours, and a partially controlled discharge-profile provides additional and unexpected discharges outside of core hours. The latter is usually unplanned and difficult to control but should be considered since such discharging occurs in reality. This project will utilise the all-hours discharge-profile only since this method is believed to provide a more realistic environment.

<table>
<tr><td></td><td colspan="24">No discharges</td></tr>
<tr><td></td><td colspan="8"></td><td colspan="12">Core Hours: discharge solution</td><td colspan="4"></td></tr>
<tr><td>Hour</td><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>17</td><td>18</td><td>19</td><td>20</td><td>21</td><td>22</td><td>23</td></tr>
</table>

Figure 9: Core hours discharge-profile

<table>
<tr><td></td><td colspan="24">Non-core hours: model prediction</td></tr>
<tr><td></td><td colspan="8"></td><td colspan="12">Core hours: discharge solution</td><td colspan="4"></td></tr>
<tr><td>Hour</td><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>17</td><td>18</td><td>19</td><td>20</td><td>21</td><td>22</td><td>23</td></tr>
</table>

Figure 10: All-hours discharge-profile

3.3.4. Models

For component 2, there are two essential models proposed, a model to predict the number of new patients for given features (i.e. hour of the day, day of the week) based upon features identified in the literature, and a model to predict a patients LOS using the same identified features, and additional features such as age and sex to increase the accuracy of the LOS model. Without the additional LOS features, a model will generalise a patients LOS without considering a patients age and sex which are expected to influence a patients time in hospital (Thomas, 2017).

A third model for predicting discharges is generated to enable comparisons between discharge solutions. This model will follow the same methodology as the admission model and exists to provide a baseline of expected discharges, essentially activity which would have occurred historically. An alternative method to generate historic discharge-profiles is the calculation of median or mean discharge-profiles by hour and day of the week, however the number of 0 discharges for a given hour is large and when initially explored, the median and mean for any specialty during any given hour was below 1 patient, and often below 0.5 patients. Whilst workarounds such as excluding 0 values to calculate the mean and median are acceptable, this could predict a higher number of discharge slots than is achievable thus it is decided to keep the prediction model in-line with the admission and LOS models for consistency. These models will henceforth be referred to as admission-model, los-model and discharge-model.

The literature provides evidence that models generated using SG methods can achieve the same performance (and in some cases improve the performance) as models generated using single estimators, and with less data. The study applied SG methods using a decision trees regressor, random forest regressor, gradient boosting regressor and linear regressor to generate independent models, followed by a second stage meta-regressor model to make a prediction of predictions. This method improved generalisation through reducing overfitting (Tugay & Oguducu, 2017) and as such, is selected to develop the admission-model, los-model and demand-model. With the aim of improving accuracy compared to the experiment by Tugay & Oguducu (2017), the linear regressor will be removed since it achieved the lowest accuracy during initial exploration, and an additional stage-1 estimator will be added based on the concept of ANN. MLP has been proven to successfully predict hourly demand for energy consumption and load forecasting with accuracy and is added to the SG model (Rodrigues , et al., 2014) (Ryu, et al., 2016) (Szoplik, 2015). Since the target attributes are converted classes with discreet values as opposed to continuous values (framing this as a classification problem), classifiers are used as the estimators of choice instead of regressors.

3.3.5. Simulation

The simulation of a specialty (component 3) for this project can be considered as a black-box (Figure 11) where patients are admitted into the specialty, utilise a bed resource, and are discharged from the specialty with little regard to the treatment pathways (there is usually one or more) or internal processes (there are usually many) which exist within the system. The purpose of this project is not to create an accurate portrayal of patient pathways, but to establish a framework which uses

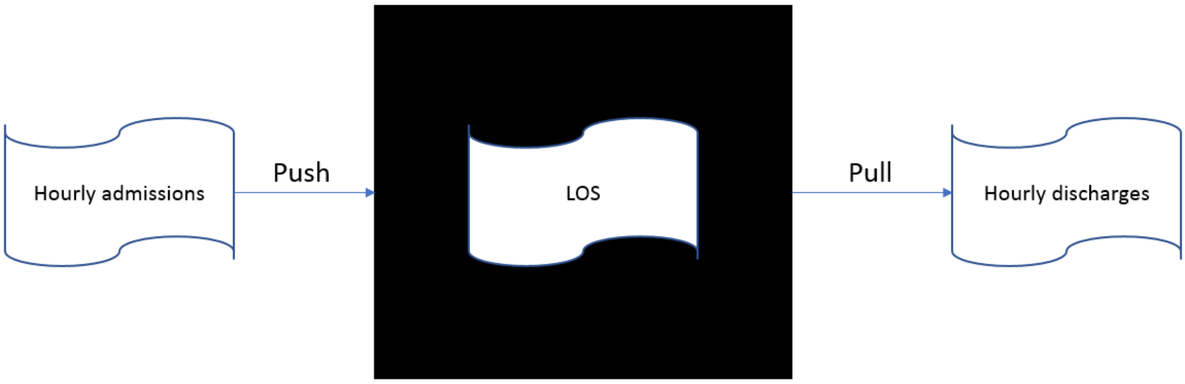

Figure 11: Simulation black box process

historic demand and activity data to generate models which represent patient flow into (and flow through) a simplified high-level black-box simulation. This can be scaled up or down in scope, or multiple black-boxes combined as part of future work so that several discharge-profiles specific to a particular element of patient pathway (ward, specialty or even waiting room) can be tested and evaluated against discharge solutions such as best practice profiles and historic profiles.

The simulation begins by predicting the number of new patients for each hour of the day which is generated using the admission-model and derived from several simulation temporal features. Each new patient requests a bed resource using a "push" method and once a bed is available, the patient will be admitted into the specialty (black-box) for a predetermined LOS (using the los-model) before being released for discharge. Even though a patient is released for discharge, the patient still occupies a bed resource until the patient is discharged. A patient can only be discharged when a discharge slot provided by a discharge-profile becomes available which is executed under a "pull" operation. Where features relating to LOS are required which can only be obtained externally to the simulation environment, the los-dataset is interrogated by the simulation to obtain said features.

A discharge-profile is dependent on one of three possible discharge solutions (component 4):

- Discharge-model solution.
- Best practice solution.
- GA solution.

This method is selected based upon research by Khanna, et al. (2016) which utilised a similar black-box method to represent an admission from ED into a specified ward without the need to specify comprehensive internal processes.

3.3.6. Genetic Algorithm

The use of GA to optimise hospitals discharges by generating a discharge-profile is one of many possible discharging solutions (component 4). A discharge-profile is a vector of discharge slots for a predefined number of hours. For this project, the use of core hours will determine the number of hours, with core hours defined as the hour 8 (8:00) through to 19 (19:59). Each discharge slot is subject to a minimum and maximum number of discharges per slot. The minimum number of hourly discharges is 0, the maximum number must be greater than 1 and cannot be larger than the number of patients waiting for discharge.

The problem can be described using the following elements:

- Let D be the number of patients waiting to be discharged.
- Let the minimum for any hourly discharge slot be 0.
- Let m be the maximum for any hourly discharge slot which is subject to $1 \leq m \leq D$ where the maximum must be an integer between 1 and the maximum number of patients awaiting discharge.
- Let h be an integer between the minimum (0) and maximum (m) number of hourly discharges. h represents the number of discharge slots available for a given hour.
- Let h_0 be the first hour of the discharge-profile (i.e. 8:00)
- Let h_n be the final hour of the discharge-profile (i.e. 19:00)
- Let H be a vector of h ($h_0, \ldots, h_n$), essentially this is a set integers forming the discharge-profile, with each h_i representing the number of slots available for a given hour.

The solution space as a chromosome can be depicted as:

$$H = [h_0 \dots, h_n]$$

subject to:

$$\{H: \sum_{h \in H} h \quad \leq D\}$$

Where:

$$\{h: 0 \leq h \leq \ m\}$$

The goals of the problem can be described with the following rules:

1. Minimise the number of lost patient hours.
    a. Patients due for discharge the following day are all released at the beginning of core hours since this is the earliest a patient could realistically be discharged.
    b. Lost patient hours is a measure of the hours between a patient being released (ready for discharge), and the hour of the day the patient is discharged.
2. Provide a staged daily discharge-profile.
    a. Manageable number of patients per hour by limiting the maximum discharges.
    b. Reduced focus on earlier discharge, try to establish some balance in the spread of discharge slots.
3. Ensure that the number of patients waiting for discharge are allocated a discharge slot during the same day.
    a. The sum of discharge slots in a discharge-profile must match the number of daily patients awaiting discharge.

These conflicting goals identify that the multi-objectives are:

1. Minimise the number of lost patient hours.
2. Minimise the deviation from a feasible staged discharge-profile.
3. Minimise the deviation from a feasible number of total daily discharge slots.

3.4. Evaluation methods

To evaluate the success of this project, two key questions need to be answered through the application of suitable tests. The first question (Table 4) evaluates the effectiveness of the framework in relation to the developed models, the second question (Table 5) evaluates the effectiveness of the framework in relation to optimising patient flow.

Table 4: Evaluation question 1

Question 1	**Is the framework effective for developing models of patient flow through a hospital setting?**
Test I	The accuracy of the models at predicting admissions, LOS and discharges against test data
Test II	The accuracy of the models at representing the flow of patient through a hospital setting

Table 5: Evaluation question 2

Question 2	**Is the framework effective for optimising patient flow through a hospital setting?**
Test III	The effectiveness of the proposed discharging solutions at reducing lost patient hours (the number of hours between the earliest hour a patient can be discharged, and the hour the patient was discharged)
Test IV	The effectiveness of the proposed discharging solutions at maximising patient flow through a hospital setting

The evaluation of test I is the mean accuracy score of the SG model against the test data (split using cross-validation), the test is applied against the admission-model, los-model and discharge-model for each specialty. This metric is selected since it provides a single value to represent the strict matching of predicted labels with the true set of labels.

The evaluation of test II is the accuracy of models when representing patient flow through a hospital setting compared to the expected flow. This is more difficult to quantify and utilises traditional data analysis techniques to generate a comparison of expectations (i.e. number of predicted annual admissions compared to the number of expected annual admissions).

The evaluation of test III is a comparison of the reduction to lost patient hours (sum of the difference in hours between the time a patient was released for discharge and the time a patient was discharged) following a simulation of 8,760 hours (1 year).

The evaluation of test IV is a comparison of the improvement to the total number of discharges following a simulation of 8,760 hours (1 year).

To satisfy both tests in question 2, an experiment is undertaken using the developed simulation environment with the following independent variables:

- Discharge solutions:
    - Discharge-model
    - Best practice
    - GA1
    - GA2
- Feasible conditions for constraints, i.e.
    - Skew between -1 and 1 (GA1)
    - Skew between 1 and 4 (GA2)

Two dependent variables are available to the simulation which will be used to evaluate the discharge solutions. These variables are:

- Total lost patient hours
- Total number of discharges

The identified control variables are:

- Random seed for probability feature selection in conjunction with the los-model (i.e. probability of age and sex)
- Random seed for converting a LOS category into a LOS value (i.e. 1 is a random number between 1 and 3).

Chapter 4. Application of chosen approach

This chapter discusses how the approach outlined in chapter 3.3 is applied using the framework components suggested in chapter 3.2, including variations to the initial approach plan. The programming language of choice to solve this problem is python 3.7 due to the availability of libraries such as scikit-learn which provides a ML framework, Distributed Evolutionary Algorithms in Python (DEAP) which provides a GA framework, and SimPy which provides a simulation framework, with each set of libraries providing detailed documentation and guides. The development environment used is PyCharm Community 2018.2. The following products are developed to satisfy the components of the framework outlined in chapter 3.2, and with each discussed in turn:

- Dataset (Component 1):
  - SQL script to extract admission-dataset
  - SQL script to extract los-dataset
  - SQL script to extract discharge-dataset
- Models (component 2):
  - Classifier.py (class which provides methods to build a model) RebuildModel.py (class to control the rebuilding of models)
  - Attributes.py (class which allows parameters to be set)
- Simulation (component 3):
  - PathwaySimulation.py (class which simulates patient flow through an element of a pathway, using the models created as part of component 2, and discharge solutions of component 4)
- Discharge solution (component 4):
  - Best practice discharge-profile (method)
  - GeneticAlgorithm.py (class which utilises several constraints to generate a good daily discharge-profile using GA)
  - Discharge-model discharge-profile (component 2)

4.1. Dataset and models

The generation of datasets and models as part of framework components 1 and 2 requires several steps. Data extraction and datasets (component 1), pre-processing data, hyper-parameter tuning with cross-validation and classifiers & stacking methods (component 2). Each step is discussed in turn, with an overview displayed in Figure 12.

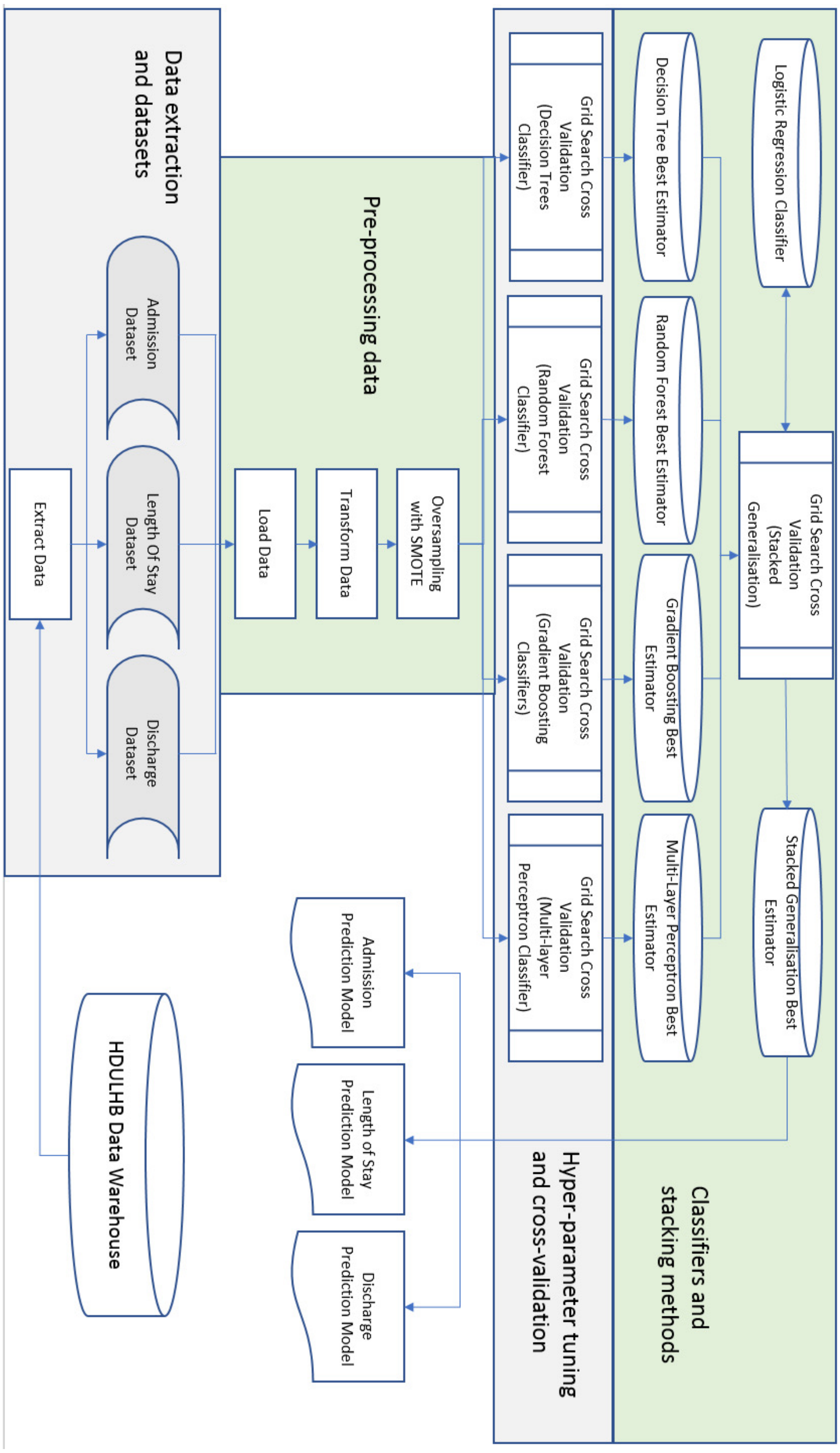

Figure 12: Process to generate optimal classification models from dataset

4.1.1. Data extraction and datasets

To generate admission-dataset, los-dataset and discharge-dataset, T-SQL scripts are used to interrogate (using SQL Server Management Studio 2018) the HDdULHB data warehouse, built on SQL Server 2008. HDdULHB provides many analytical views of data based on key patient and organisational themes. The data extraction stage is subject to the data extraction methods and conversions mentioned in chapter 3.3.2, and relates only to the ASDS view. A full list of data attributes for each dataset can be found in Appendix II.

Each specialty (ENT, T&O and URO) requires an admission-dataset, los-dataset and discharge-dataset which are prefixed with the specialty (i.e. ent-admission-dataset) and saved as a .csv file with column headers included. Once generated, the dataset is loaded using the RebuildModels.py class which calls upon the Classifier.py class (utilising a csv reader method) and begins the generation of a new model.

4.1.2. Pre-processing data

The extracted datasets are the first elements loaded to a Python application. The classifier.py class provides methods to build a model, load and save models, import data, making unseen data predictions and performing hyper-parameter and cross-validation using GridSearchCV. Two transformers (RobustScaler and MinMaxScaler) are used to pre-process data, facilitated through the scikit-learn pre-processing package which contains several utility functions and transformer classes (Raghav, et al., 2017).

RobustScaler removes the median of a given feature and scales the data according to the inter-quartile range. This function is applied to minimise the impact of outliers on several dataset features such as Age and LOSGroup (Raghav, et al., 2017) (Cao, et al., 2016). MinMaxScaler rescales the data so that all features are in the range [0, 1]. This function is applied since some ML classifiers (such as MLP) perform well with features within the range [0, 1] (Raghav, et al., 2017).

A backbone concept to the data flow (although not integral to all methods) is a pipeline which is a sequential order of data transformations and estimators that are applied in turn. Pipelines can be simultaneously cross-validated and a data fit (where an estimator is fit to the data in order to predict classes) is only required and performed by the final estimator.

The project approach outlines no plan to use sampling techniques, however following development of several models which provide acceptable accuracy (see chapter 5.1), yet for any given set of features predicts 0 patient admissions and discharges, further investigation into data imbalances is prompted.

Typical imbalances between classes with label 0 and all other class labels varied between 1:1 and 10:1, however when considering the imbalance between class labels [0, 1] and [2, n], imbalances varied between 2:1 and 63:1. All imbalances are available in Table 6. Note that this is an imbalance between groups of classes, if the class with smallest representation is compared to the class with largest, the imbalances are even greater.

Table 6: Data imbalances between majority and minority classes

Specialty	Dataset	Imbalance [0]:[1,n]	Imbalance [0,1]:[2,n]
ENT	Admission	9:1	45:1
ENT	Discharge	9:1	45:1
ENT	LOS	10:1	30:1
T&O	Admission	6:1	63:1
T&O	Discharge	6:1	62:1
T&O	LOS	1:1	2:1
URO	Admission	8:1	39:1
URO	Discharge	7:1	45:1
URO	LOS	3:1	8:1

An investigation into sampling techniques found potential solutions to correcting the imbalance, with imblearn library offering the over_sampling package which contains functions to perform oversampling, undersampling and SMOTE (Lemaitre, et al., 2016). Each function is applied to the data, and model accuracy compared, finding that contrary to the literature which states that undersampling often scores highest (Dittman, et al., 2014) (Mishra, 2017), undersampling performed worst with a 41-72% (depending on the dataset in question) drop in model accuracy against test data. Oversampling performed better with a 15-26% decrease to model accuracy, however SMOTE performed best with a decrease of approximately 2-12%.

SMOTE is selected as the sampling method, although an issue identified as part of the development process is that the scikit-learn pipeline applies SMOTE prior to a cross-validation split which means that oversampled values exist in the test data, providing false accuracy. While this means that SMOTE may provide improved performance against test data (potentially resulting in overfitting), a negative impact is increased likelihood of poor performance against new unseen data.

To mitigate this issue, splitting training and test data occur prior to applying SMOTE by loading SMOTE (and the transformers) to the imblearn pipeline which splits training and test data prior to the application of pre-processing functions and generally handles samples correctly, with an added benefit of operating effectively in conjunction with GridSearchCV (Lemaitre, et al., 2017) (Aridas & Lemaitre , 2016).

4.1.3. Classifiers and stacking methods

Based upon the method used by Tugay & Oguducu (2017) to predict hourly demand of an online e-commerce company, and evidence discussed in the literature suggesting SG can improve model generalisation through reduced overfitting (which is useful for generalised simulations where general flow is just as appropriate as accurate predictions), SG is selected to develop admission-model, los-model and discharge-model. This project applies some of the methods outlined in that study, with some modification i.e. whilst most class labels in admission-dataset and discharge-dataset map directly to a continuous value (i.e. class 1 is 1 patient, class 2 is 2 patient), some classes did not provide sufficient representation for effective cross-validation, thus the data conversions outlined in chapter 3.3.2 are applied, transforming values into discreet class labels. Additionally, all the LOS target classes are discreet class labels, thus it is more appropriate to use classifiers as opposed to regressors. Another modification aiming to improve model accuracy is the removal of the accurate estimator (linear regression) which is replaced by MLP since it is proven to perform well when modelling demand in the field of energy production and consumption.

The meta-classifier used to create an SG solution is StackingCVClassifier which is part of the mlxtend.classifier library (Raschka, 2018). StackingCVClassifier essentially works as a container for the stage-1 and stage-2 classifiers and applies cross-validation as part of the process. Table 7 provides a list of classifiers used to form an SG method, including the stage at which the classifier resides, and its reason for selection as part of this project.

Table 7: Classifiers included in SG model

Classifier	Stage	Reason
DecisionTreeClassifier (DT)	Stage-1	Used in another study.
RandomForestClassifier (RF)	Stage-1	Used in another study.
GradientBoostingClassifier (GB)	Stage-1	Used in another study.
MLPClassifier (MLP)	Stage-1	Used for demand modelling in other fields.
LogisticRegressionClassifier (LR)	Stage-2	Classification alternative to linear regression which was used in another study as a stage-2 generaliser.

Each stage-1 classifier is tuned independently since there is an issue with tuning the hyper-parameters of stage-1 estimators where they are already part of a SG model. The author of this project was unable to resolve the issue within reasonable time, thus GridSearchCV is performed on each classifier which is added to the SG model once a best estimator is identified. GridSearchCV is further applied to tune the hyper-parameters of a stage-2 meta-classifier, before being added to the data pipeline for model fitting, with the finished model saved as a pickle (.pkl) file.

4.1.4. Hyper-parameter tuning with cross-validation

Hyper-parameter tuning is a standard approach and is often necessary for some problem domains to improve model performance although results can vary greatly from one domain to another (Claesen, et al., 2017). Since initial model accuracy is lower than anticipated, particularly due to a reduction of accuracy post-sampling, GridSearchCV (using the scikit-learn model_selection package) is utilised which allows the specification and selection of hyper-parameters, with additional functionality of performing cross-validation against said hyper-parameters while retaining the best estimator (defined as the estimator with the most accurate cross-validation score). Cross-validation is set at K-Folds=5 which repeats stratified K-Folds five times and is selected since a lower value provided lower accuracy scores, and a higher value often provided insufficient training data class representation.

A baseline of model accuracy for each specialty is gathered using the default settings provided on the scikit-learn website (scikit-learn developers, 2017). GridSearchCV is applied using several sensible alternative parameters (i.e. for hyper-parameters with an initial rate of 0.5, the first round of tuning utilised bounds of 0.1 and 0.9 since this is a sensible range) with incremental adjustments for each consecutive cycle of tuning i.e. if initial rate is 0.5, and 0.1 provided improved accuracy compared to 0.5 and 0.9, the tuning search space is adjusted by increments of 0.1, with the new search parameters of 0.2, 0.4 and 0.8. Tuning cycles continue until optimal hyper-parameters are identified. Where there is less than 1% variance between specialty model performance for specific hyper-parameters (i.e. min_samples_split for admission-model varies by less than 1% for ENT, T&O and URO), the best performing hyper-parameter value is selected to reduce future tuning cycle computational demands. Likewise, once final tuning is complete, where there is less than 1% variance between admission-model, los-model and discharge-model, the most accurate hyper-parameters are selected to reduce the time to generate future models. The use of gini and entropy has a variance of up to 5% when applied to any model, thus clf__criterion of gini and entropy for DT and RF is always performed using GridSearchCV to improve accuracy. The following hyper-parameters are tuned:

Table 8: Hyper-parameters

Hyper-parameter	Classifier	Initial value	Increment rate	Final value
criterion	DT	gini	gini/entropy	gini/entropy
min_samples_split	DT, RF, GB	2	1	2
min_samples_leaf	DT, RF, GB	1	1	1
max_features	DT, RF, GB	auto	auto/log2/10	auto
learning_rate	GB	0.1	0.1	0.3
hidden_layers	MLP	100,	100,	1000,1000,1000
max_iter	MLP	100	100	200
momentum	MLP	0.1	0.1	0.1
warm_start	MLP	false	true/false	false
early_stopping	MLP	false	true/false	true
n_estimators	RF, GB	10/100	10/50	100/200
use_probas	SG	false	true/false	true
use_clones	SG	true	true/false	false
shuffle	SG	true	true/false	true
stratify	SG	true	true/false	false

4.2. Discharge solutions

Discharge solutions generate and provide a discharge-profile (Figure 13) for the purpose of managing outbound flow, forming component 4 of the framework. The three discharge solutions are best practice discharge solution (BPDS), discharge-model discharge solution (DMDS) and GA discharge solution (GADS), with only BPDS and GADS discussed in this chapter since DSDS is generated as part of chapter 4.1.

	Non-core hours																							
									Core hours															
Hour	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Slots	1	0	0	0	1	0	0	0	2	1	3	0	0	0	1	0	0	0	0	1	1	0	0	0

Figure 13: Example discharge-profile

4.2.1. Best practice

The best practice discharge solution is a method provided by the simulation system based upon evidence that 80% of patient discharges occurring between 8am and 11am, and 100% of patients discharged before 12pm is an optimal discharge-profile.

The method calculates 80% of patients waiting for discharge by multiplying the number waiting for discharge by 0.8 to the nearest whole number, divides the number by 3 to the nearest whole number, this value is used as the 8am (slot1), 9am (slot2) and 10am (slot3) discharge slots. All remaining patients form the 11am discharge slot (slot4), which is calculated by removing the value calculated for 80% of patients from the number awaiting discharge.

4.2.2. Genetic algorithm

DEAP is an evolutionary framework which encourages development of suitable evaluation functions based on the specific requirements of the problem domain. DEAP allows prototypes and ideas to be developed quickly and provides users with types (such as individuals and fitness – although user defined types can be constructed) and a toolbox of operators and algorithms (DEAP Project, 2018).

The purpose of the GADS is to produce a discharge-profile subject to several constraints outlined in chapter 3.3.6. This is a multi-objective requirement, balancing the number of lost patient hours with later-day discharges. Two additional objectives provide further solution quality: the sum of discharge slots in discharge-profile, and the maximum value of a discharge slot. Any violation of the former may constrict patient flow by generating a discharge-profile with fewer slots than patients awaiting discharge. A violation of the latter can produce an unachievable discharge profile which is technically balanced, but not optimally balanced. (i.e., 6,0,0,0,0,6 may provide similar balance to 2,2,2,2,2,2, but is more difficult to achieve by the service). There is merit to using alternative balancing techniques (i.e. discharge slots based upon known ward rounds, where discharge decisions are made) and suitable constraints can be developed for use with this framework, however this project considers only discharge-profile distribution skew and limiting upper-bounds of discharge slots to balance the profile for the purpose of keeping the framework simple. The following methods calculate an objective value, with notational definitions available in chapter 3.3.6.

x_1 is the sum all discharge slots available in a discharge-profile and forms the objective value of constraint C1:

$$x_1 = \sum_{h \in H} h$$

x_2 is the maximum discharge slots for a single hour of a discharge-profile (for this project, ¼ of patients awaiting discharge (D) is the limit) and forms the objective value of constraint C2.

$$x_2 = \frac{D}{m}$$

where:

$$m = 4$$

x_3 is the sum of lost patients hours (the number of hours between the patient being ready for discharge and the hour the patient is discharged) and forms the objective value of constraint C3. For patients discharged before 8am, the number of hours between midnight and 8am is removed, while patients discharged on or after 8am are deducted eight-hours. This ensures that discharge-profiles are only penalised for lost hours after the beginning of a discharging day (i.e. there may not be suitable patient transport at 5am).

$$x_3 = \sum_{n=0}^{23} h_n \quad \begin{cases} h_n - 8 & n \geq 8 \\ h_n - n & n < 8 \end{cases}$$

x_4 is the distribution skew of the discharge-profile and forms the objective value of constraint C4.1 and C4.2. Two feasibility ranges are used to compare simulations, skew1 forming the GADS1 discharge solution, and skew2 forming the GADS2 discharge solution.

$$x_4 = Skew(H)$$

Table 9 provides the constraints C1-4.2, objectives (x) which are minimised, feasibility, delta ($\triangle$), centre for distance measure x_0 and penalty (quadratic distance) applied. C3 does not utilise feasibility thus the centre is 0, C3 simply requires minimisation as there is no definition of acceptable lost hours. Additionally, C4.1 and C4.2 distance from centre is multiplied by 10 due to skew values being in range 0-1 which is too small compared to other objective values which provide larger ranges.

Table 9: Objectives, feasibility and penalty calculations

ID	Value $\boldsymbol{x}$	Feasibility	Delta $\triangle$	Centre $\boldsymbol{x_0}$	Penalty
C1	x_1	$D \leq x_1 \leq (D+2)$	100	$\frac{D+(D+2)}{2}$	$\triangle + (x - x_0)^2$
C2	x_2	$(m-1) \leq x_2 \leq (m+1)$	10	$\frac{(m-1)+(m+1)}{2}$	$\triangle + (x - x_0)^2$
C3	x_3	N/A	10	0	$\triangle + (x - x_0)^2$
C4.1	x_4	$-1 < x < 1$	10	$\frac{-0.5+0.5)}{2}$	$\triangle + (10 \cdot (x - x_0))^2$
C4.2	x_4	$1 < x < 4$	10	$\frac{1+4}{2}$	$\triangle + (10 \cdot (x - x_0))^2$

An example of the methods used to evaluate feasibility of x_1, x_2, x_3 and x_4 and calculate a distance penalty where objective values are deemed unfeasible is provided in Figure 14 where the method sets and evaluates the feasibility bounds.

```
# method to check whether skew is within the required constraints
@staticmethod
def feasible_skew(skew_value):
    # set the min and max bounds of feasible solution
    min_bound = 1
    max_bound = 4
    # check if the skew value if in the feasible range
    if min_bound < skew_value < max_bound:
        # if so, return low score of 0
        return 0
    # if not, calculate the centre point for the distance penalty
    centre = (min_bound + max_bound) / 2
    # return a quadratic distance penalty based on the distance between the skew value and the centre
    # multiply the difference between skew and centre by 10 otherwise the values are too low compared to other
    # constraints.
    return _a.delta_penalty + (((skew_value - centre) * 10) ** 2)
```

Figure 14: Constraint C4.2

If the value is within feasible bounds, a perfect score of 0 is returned to the parent function, otherwise a quadratic distance function is applied. An important aspect of a quadratic distance function is the delta penalty. A delta penalty of 10 is applied for all constraints with exception to C1 which experiences convergence issues where many discharge-profiles provided an infeasible number of discharge slots per discharge-profile. This issue is reduced by applying a delta penalty of 100. By applying a higher delta, C1 mimics the hard-constraints discussed in the literature, with C2-C4.2 acting as soft-constraints which provides a higher probability that C1 will be satisfied (Puente, et al., 2009).

The GA is called by simulation using the GeneticAlgorithm.py class and provides a starting population of 2,000 solutions with a limit of 200 generations before the search terminates, selecting the best scoring discharge-profile. The solution also terminates if no scoring improvements are made for 30 generations. This configuration provides a good balance between discharge-profile generation for use with simulation, computational power requirements and suitability of discharge-profiles, although the values can be increased where necessary to provide further optimised solutions. All weights for the four objectives of the multi-objective function are assigned a value of -1.0 which signifies that minimisation objectives are used. The evolution strategy applies the evolutionary concepts of crossover (Table 10), mutation (Table 11) and selection (Table 12) to evolve good solutions.

Table 10: Crossover strategy

Crossover probability (cxpb)	0.7
Operator	cxUniform
Description and parameters	Attributes are swapped through a uniform crossover Indpb = 0.2

Table 11: Mutation strategy

Mutation probability (mutpb)	0.3
Mutation operator	mutUniformInt
Mutation operator description and parameters	Attributes are swapped (mutated) by replacing attributes using a random integer between the lower (low) and upper (up) bounds of a set range. Indpb = 0.5 Low = Minimum hourly discharges (0) Up = Maximum hourly discharges (x_2)

Table 12: Selection strategy

Selection operator	selBest
Selection operator description and parameters	Selects the k best individuals based on an evaluation of fitness.

Figure 15 demonstrates the evolutionary process for crossover, mutation and selection. In the code, a group of offspring is generated from a population using the selection strategy. Two children are extracted from the offspring and assessed against a randomly generated number according to a predefined crossover probability of 0.7 (cxpb). Where the number is less than cxpb, a mating strategy is applied, and the fitness values are removed since they are recalculated later. Each child in offspring is then assessed for mutation by comparing a random number and the predefined mutation probability of 0.3 (mutpb). Where the number is less than mutpb, a mutation strategy is applied, and the fitness values are removed. Individuals are then evaluated for fitness (using the evaluation function) and the entire population is replaced with a new generation. The above probability values are selected since they seemed sensible, are the default values provided, and produced good solutions.

```python
# select the next generation individuals
offspring = self.toolbox.select(pop, len(pop))
# clone the selected individuals
offspring = list(map(self.toolbox.clone, offspring))
# apply crossover and mutation on the offspring
for child1, child2 in zip(offspring[::2], offspring[1::2]):
    # cross two individuals with probability CXPB
    if random.random() < _a.cxpb:
        self.toolbox.mate(child1, child2)
        # fitness values of the children
        # must be recalculated later
        del child1.fitness.values
        del child2.fitness.values
# apply mutation
for mutant in offspring:
    # mutate an individual with probability MUTPB
    if random.random() < _a.mutpb:
        self.toolbox.mutate(mutant)
        del mutant.fitness.values
# Evaluate the individuals with an invalid fitness
invalid_ind = [ind for ind in offspring if not ind.fitness.valid]
fitnesses = map(self.toolbox.evaluate, invalid_ind)
for ind, fit in zip(invalid_ind, fitnesses):
    ind.fitness.values = fit
# The population is entirely replaced by the offspring
pop[:] = offspring
# select best individual
best_ind = tools.selBest(pop, 1)[0]
# compare best individual with all time best, replace if better scoring solution
if sum(best_ind.fitness.values) < sum(all_time_best.fitness.values):
    all_time_best = best_ind
    g_i = 0
else:
    g_i += 1
```

Figure 15: Evolution

4.3. Simulation

The simulation generated for this project represents a model of patient flow through a hospital setting (essentially bringing components 2 and 4 together) and forms component 3 of the framework. Figure 16 provides an overview of the simulation flow process, with SimPy selected as the development framework. SimPy is a discreet-event, process-based simulation framework which uses Python generator functions to model objects, interactions and resources (Team SimPy, 2018).

A simulation begins by loading the PathwaySimulation.py class developed for this framework. SimPy is controlled by an environment process which manages time and scheduling, and processes events. Once an environment is executed, it simulates the environment until a stop criteria is met, usually a set number of clock intervals. For this project, the clock interval of 1 represents each hour of the day (0-23), with the simulation terminating following 8,760 cycles (1-year).

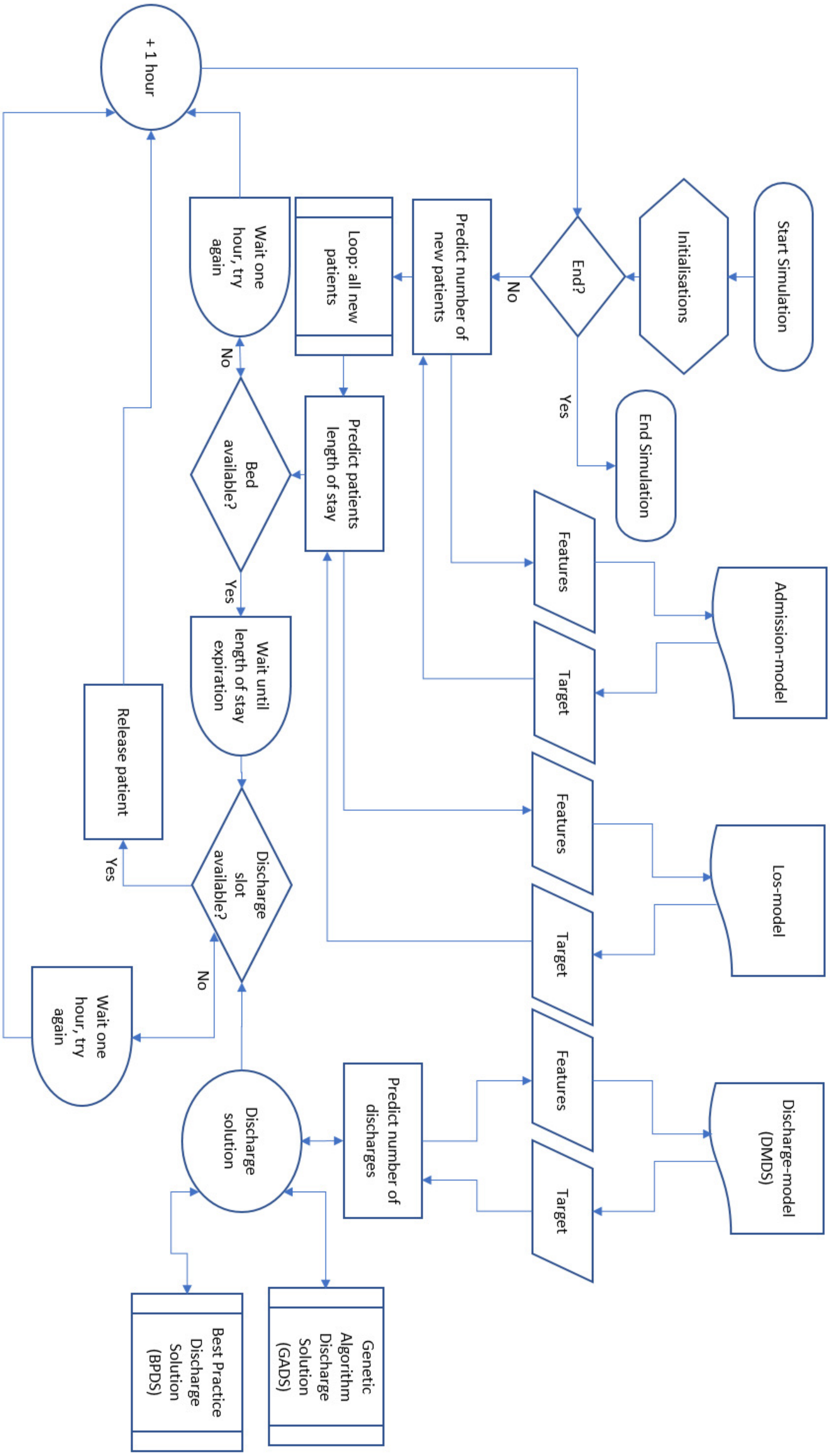

Figure 16: Simulation process chart

A calendar method is used to keep track of temporal related features, with the first processes executing at midnight on 1st April 2018. The method generates a binarized calendar which provides temporal features to the models outlined in chapter 4.1. Additional features such as age and sex (required by los-model) are obtained using a method to extract a distribution profile from los-dataset. The distribution is calculated by extracting the age attribute from los-dataset and grouping the attribute by age value. The count of age attributes are added to a list and a sum of all counts is calculated. Each age count is then divided by the sum to calculate the age probability and added to a new list. The distribution is further added to a random.choice function which calculates a patients age and sex based upon probability. The same method is used to calculate the probability of males and females in los-dataset.

For each hour of the day, simulation uses admission-model to predict the number of new patients, and for each new patient, simulation uses los-model to predict a patients LOS category which is then converted into days using a random number generator based upon the conversions outlined in chapter 3.3.2, i.e. LOS=1 becomes a random number between 1 and 3. The value is then multiplied by 24 to calculate the number of hours, however since the number of patients ready for discharge should be known at midnight (for following day discharging), patients are released at 11pm the night before by removing the time difference (hours) between a patients admission and midnight. The idea is that a following day discharge-profile is generated at midnight and with sufficient planning, the service would be aware that a patient may be available for following day discharge.

Once these elements are calculated, each new patient needs a resource, of which there are three types available to SimPy, regular, priority and pre-emptive. A regular resource can be requested by a process, has a first-in-first-out (FIFO) queue, has a finite capacity and can be released by a process. A priority resource offers additional functionality of prioritisation, where pending requests are sorted in priority ascending order, and a pre-emptive resource uses pre-emption to override lower priority processes already occupying a resource – ensuring that the highest priority processes are assigned to a resource immediately (subject to capacity) (Team SimPy, 2017). The two resources used are regular and pre-emptive.

Bed capacity is a regular resource and available to processes (patients) on a FIFO basis. If a bed is unavailable the process yields a timeout for 1-hour before trying again. For improved simulations, a priority resource could be used to accept patients based on priority (i.e. emergency vs. elective patients), however for this basis of this framework and simplification of the flow, no prioritisation is used.

Once a patients LOS has expired, the patient is ready for discharge, although the patient still occupies a bed until discharged. A discharge can only be facilitated though an available discharge slot, effectively acting as a pull mechanism to pull a patient out of a bed. The number of available discharge slots is a pre-emptive resource, the reason for this is that a regular resource cannot change capacity once initialised which is problematic to the varying discharge slots per hour, therefore a valid workaround is to set a larger than required pre-emptive capacity (i.e. the number of beds) and block all the capacity slots with low priority processes. Then for each hour, override (pre-empt) the resource with the number of discharge slots provided by the discharge-profile using processes with a higher priority than the blocking processes. All processes are released and renewed each hour of the day. A discharge slot is provided through a discharge-profile, which is generated by one of the discharge solutions outlined in chapter 4.2. Depending on the solution requested, simulation will call upon the relevant method to generate the discharge-profile.

A series of five simulations are executed for each discharge solution, and also for each specialty. There are three specialties and four discharge solutions, thus sixty simulations are provided. The reason for this is that whilst the number of admissions will remain constant during each simulation by using a defined random seed, as will the discharge-model predictions and most SimPy processes, the los-model predictions will vary since the numpy.random.seed() method using probability to generate age and sex is not deterministic where there are several threads (i.e. patient processes). Workarounds exist for this problem; however, it was decided to maintain los-model (which scored lower accuracy as discussed in chapter 5) as non-deterministic to introduce some random variation to the simulation black-box. A median score is calculated for each specialty discharge solution.

Chapter 5. Analysis and limitations

This chapter provides analysis of the results, and the projects limitations. The project is analysed relative to the effectiveness of the framework for developing models representing patient flow, and effectiveness of optimising patient flow. Test I and II (identified in chapter 3.4) are applied to the suitability of the frameworks models, while test III and IV are applied to evaluate discharge-model discharge solution (DMDS), GA discharge solutions (GADS1 and GADS2) and best practice discharge solution (BPDS). Each discharge solution is simulated five times per specialty (sixty simulations total) to provide a range of results.

5.1. Analysis of results (test I)

Test I establishes the accuracy of the models against test data using a scoring accuracy generated during component 2 of this framework. The scoring accuracy results are compared in Table 13 for the following methods:

- untuned hyper-parameters with no transformations applied (S1)
- untuned hyper-parameters with transformers and SMOTE applied (S2)
- tuned hyper-parameters with transformers and SMOTE applied (S3)

The length of time (hours) required to generate each model is also included in the table since this provides awareness of the limitations to further-tuning based upon the computational power available. While this element does not relate directly to Test I, it does relate to question 1 which asks whether the framework is effective.

Table 13: Model tuning results

Specialty	**Dataset**	**S1**		**S2**		**S3**	
		Accuracy	Hours	Accuracy	Hours	Accuracy	Hours
ENT	Admission	0.900	0.4	0.805	4.4	0.855	32.8
ENT	LOS	0.915	0.1	0.887	1.6	0.901	2.7
ENT	Discharge	0.895	0.2	0.789	5.5	0.792	18.8
T&O	Admission	0.855	0.3	0.733	2.5	0.763	14.0
T&O	LOS	0.528	0.1	0.479	2.0	0.484	7.6
T&O	Discharge	0.876	0.3	0.768	5.6	0.772	15.6
URO	Admission	0.890	0.4	0.815	2.9	0.854	16.0
URO	LOS	0.743	0.1	0.647	0.4	0.693	1.0
URO	Discharge	0.886	0.4	0.783	2.6	0.787	17.2

Seven out of nine S3 models predict against test data with an accuracy of greater than 75%, with three models greater than 85% which for the field of healthcare capacity and demand predictions, is respectable. The results from test I suggest that the framework is effective for developing models of patient flow, with only los-model for T&O and URO deemed unsatisfactory with a caveat that additional patient attributes may improve los-models which can be facilitated by this framework.

Table 14 provides the accuracy change to S2 compared with S1 when S2 is applied, followed by the accuracy change to S3 compared with S2 when S3 is applied and finally the accuracy change between S1 and S3.

Table 14: Changes to accuracy following transformation and tuning

Specialty	Dataset	S1 to S2 accuracy change	S2 to S3 accuracy change	S1 to S3 accuracy change
ENT	Admission	-0.095	0.050	-0.045
ENT	LOS	-0.028	0.014	-0.014
ENT	Discharge	-0.106	0.003	-0.103
T&O	Admission	-0.122	0.030	-0.092
T&O	LOS	-0.049	0.005	-0.044
T&O	Discharge	-0.108	0.004	-0.104
URO	Admission	-0.075	0.039	-0.036
URO	LOS	-0.096	0.046	-0.050
URO	Discharge	-0.103	0.004	-0.099

It is evident that the S1 models obtain the highest accuracy scores, however several models fail to predict some (or all) minority classes which rendered the real-world application of S1 null. Applying S2 results in accuracy declines, typically by 3-12%, although simulation results observed improved minority class representation making the models more usable by this framework. This means that S2 is essential for the effectiveness of this framework. Minor improvements to accuracy scores are gained by applying S3, although it is difficult to establish whether a typical improvement of around 1-5% is worth the additional hours (typically 15-30) to perform S3. This renders the hyper-parameter tuning aspect of component 2 as an optional step in this framework.

Table 15 provides the S3 scoring accuracy of each individual classifier and specialty/model combination, with each best scoring models highlighted in bold.

Table 15: Final tuned classifier accuracy

Dataset	Classifier	ENT	T&O	URO
Admission	DT	0.826	0.746	0.827
Admission	RF	0.822	0.739	0.826
Admission	GB	**0.864**	**0.773**	**0.859**
Admission	MLP	0.803	0.723	0.818
Admission	SG	0.855	0.763	0.854
LOS	DT	0.831	0.381	0.595
LOS	RF	0.896	0.462	0.670
LOS	GB	0.878	0.455	0.674
LOS	MLP	0.850	0.422	0.611
LOS	SG	**0.901**	**0.484**	**0.693**
Discharge	DT	0.795	**0.775**	0.789
Discharge	RF	**0.795**	0.772	0.786
Discharge	GB	0.788	0.722	0.786
Discharge	MLP	0.771	0.753	0.774
Discharge	SG	0.792	0.772	**0.787**

While S3 tuning offered accuracy increases of up to 15% for some individual models (particularly GB and MLP), the increases incurred by the final SG models are not as large. It is possible that lower scoring models (such as DT for los-models) may contribute to lower SG scores and this concept should be explored. Further tuning is still possible, i.e. a test tuning-cycle for ent-admission-model using MLP with hidden_layers set at 5000,5000,5000 neurons, and GB n_estimators set at 1500 found increased accuracy of 2% to MLP, 1% to GB and 1% to SG, however the model requires more than 4-days to generate with the available computing power thus are omitted as hyper-parameter options in this project. SG scored higher accuracy than any individual generaliser for each specialty los-model, GB scored best for each admission-model while discharge-model best estimators are mixed, although SG often scored within the top-2 best estimators for all combinations suggesting that the method is suitable for this framework.

5.2. Analysis of results (test II)

A key measure to evaluate whether the models effectively represent patient flow is the total number of patients admitted and discharged through each specialty when applying all three developed models (admission-model, los-model and discharge-model) in a simulation.

Table 16: Comparison of historical volume vs. simulated volume

Specialty	Type	Historical average (annual)	Simulated (annual)	Simulated vs. historical
ENT	Admission	1,827	2,285	+25%
ENT	Discharge	1,825	1,942	+6.4%
T&O	Admission	2,365	2,542	+7.5%
T&O	Discharge	2,355	2,320	-1.5%
URO	Admission	1,453	1,765	+21%
URO	Discharge	1,441	1,358	-6%

The five-year average historical volume, and the simulated volume of patient admissions and discharges are found in Table 16. The results find some variance between the historical and simulated predictions, with all but two models within a 10% tolerance which suggests that components 1 and 2 of this framework can be applied with component 3 to effectively model patient flow. A 21% and 25% variance is also arguably reasonable for the following reasons:

1. Demand is set to increase, thus higher volume models can help future planning
2. This framework is more concerned with the optimisation of flow using representative models as opposed to highly accurate models
3. Existing methods to calculate demand will include some variance

While these reasons do not prove that the framework is effective, they suggest that the higher variances of 21% and 25% do not render the framework ineffective.

To further investigate model representation, analysis of historical profiles are compared to predicted profiles. Prior to S2 in component 2, predictions for ENT did not represent a typical admission (Figure 17) or discharge (Figure 18) profile. Note that results for T&O and URO profiles are not included since S1 models failed to predict any admissions or discharges.

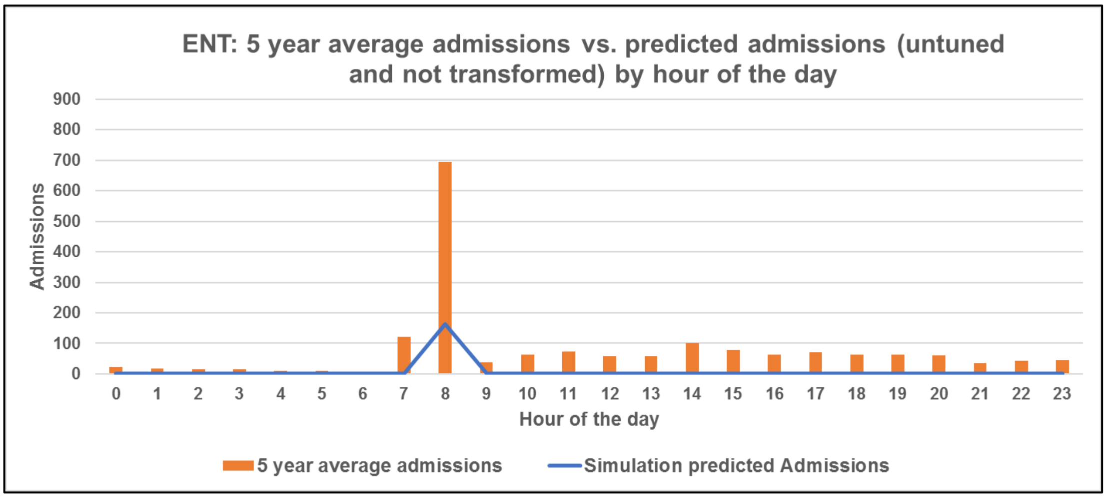

Figure 17: Untuned ENT hourly admissions

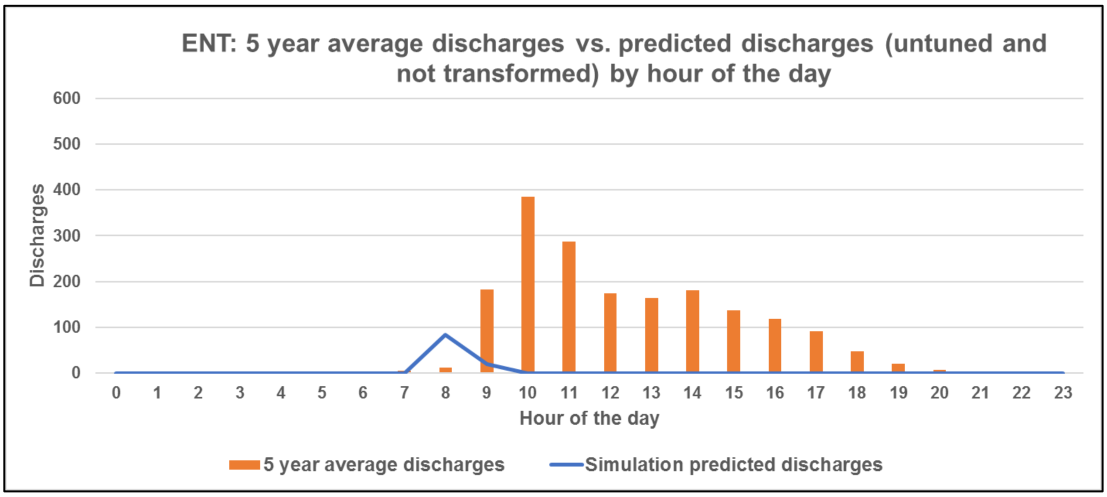

Figure 18: Untuned ENT hourly discharges

Following S3, it becomes clear that the admission (Figure 19) and discharge (Figure 20) profiles for ENT provide a better representation of historical data, although whilst the general prediction patterns appear representative, the models often predict too few or too many patients during some hours. It is possible that this is influenced by the application of SMOTE and tuning the SMOTE bias may further improve the models.

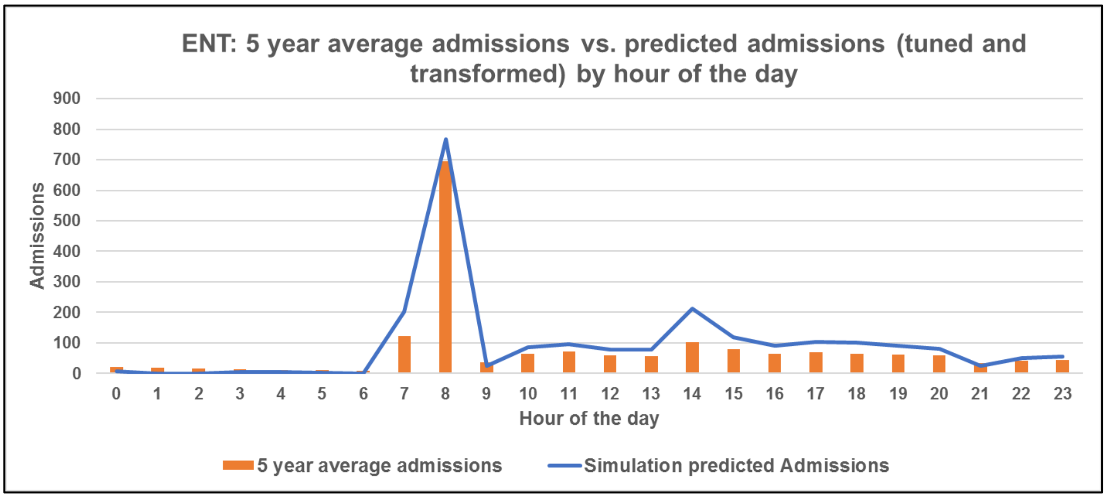

Figure 19: Tuned ENT hourly admissions

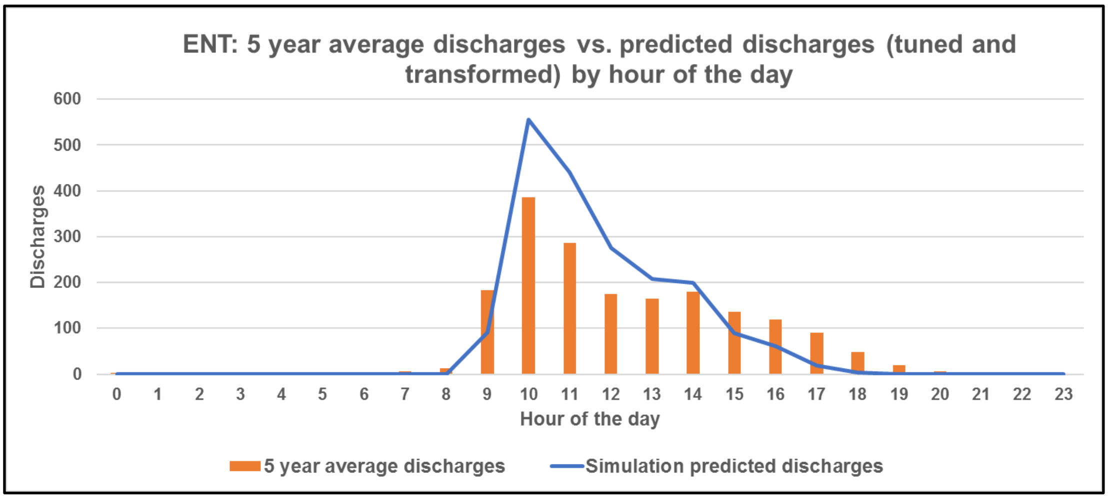

Figure 20: Tuned ENT hourly discharges

The results of T&O admissions in Figure 21 suggest that predicted admissions made by admission-model reasonably represents historical data, although improvements are needed to the accuracy of predictions between 1am and 10am (where there are no predicted admissions) and later-day (where there are too many predicted). This may explain a discrepancy of historical annual admissions against predicted annual admissions and could be improved by further tuning the models, adding new features which represent the timeframes in question, or tuning SMOTE. The discharge profile (Figure 22) occupied similar concerns, although unlike previous predictions for ENT and T&O, the discharge-model for T&O fails to identify the correct peak.

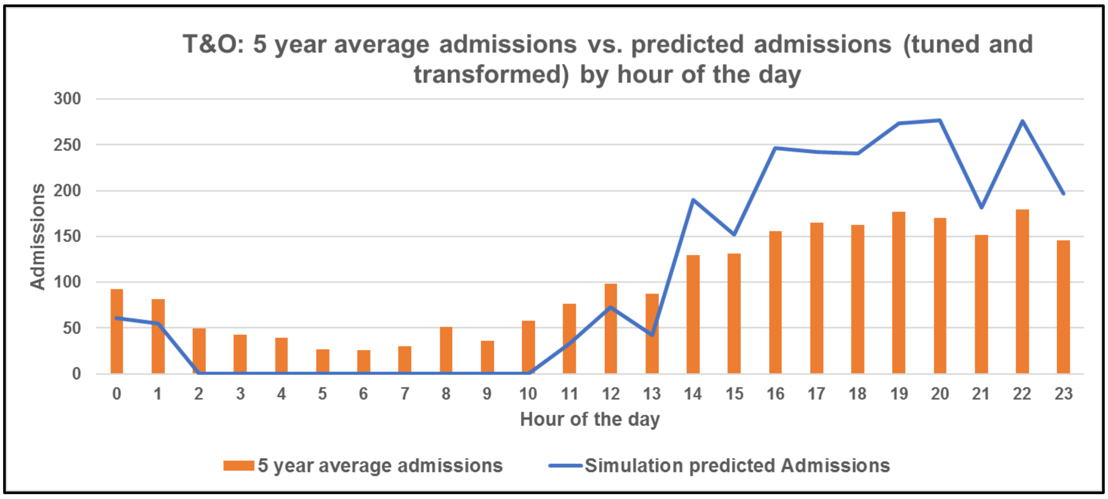

Figure 21: Tuned T&O hourly admissions

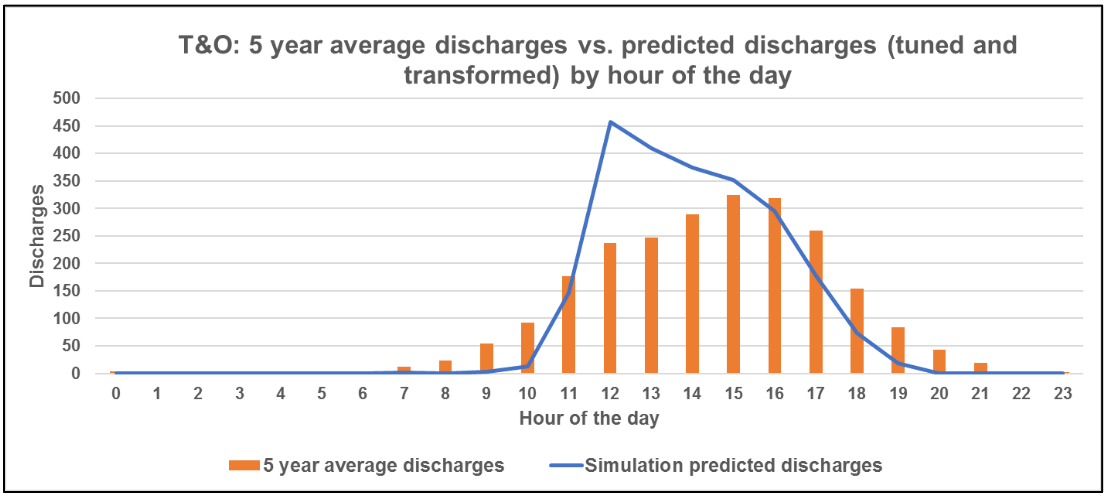

Figure 22: Tuned T&O hourly discharges

The results for hourly URO profiles display similar results and are omitted from this report. A final measure of model representation is average daily admission profiles which are used to evaluate trends, with a 10-day rolling average applied to historical and simulated results to smooth the trendlines. The ENT admission profile in Figure 23 is relatively representative, although interestingly there is a disconnect during known healthcare events. Admission-model fails to match a decrease of admissions during early-summer (July-August), and during the initial winter phase (November-December), the model also predicts larger volumes of admissions during January and March.

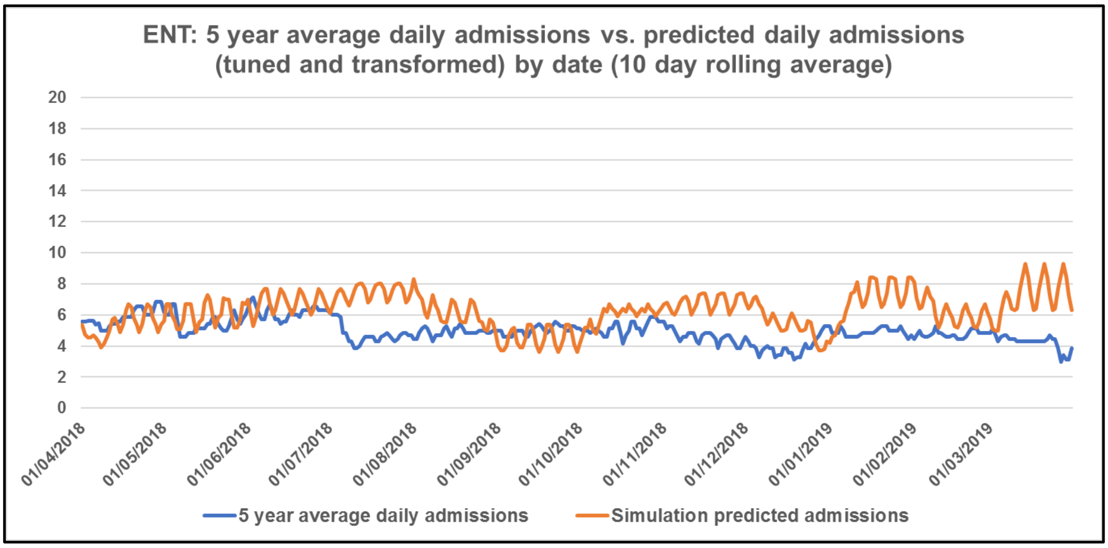

Figure 23: Tuned ENT daily discharges

The T&O admission profile (Figure 24) is more reflective of historical trends than ENT, although predicts higher admission rates during June and September, and generates several sharp declines to admissions (i.e. May, July, October and January) which are not reflective of historical data.

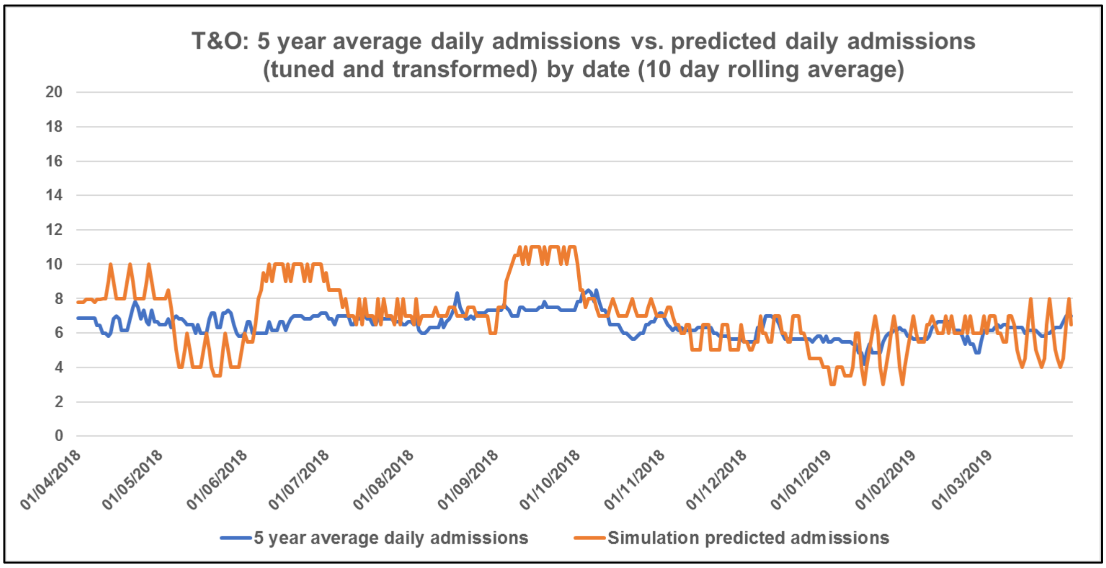

Figure 24: Tuned T&O daily admissions

The reasons for these inconsistencies is unknown, although a possible explanation is outlined in the literature and relates to the multiple processes found within the data. i.e. the data includes both elective and emergency patients, and during some annual events (such as Christmas, summer holiday, end of financial year), the number of elective patients may be increased or constrained to cope with varying demand and capacity. It is possible that admission-model, los-model or discharge-model may underfit the data since there are no attributes used with component 1 which reflect these sub-processes. Interestingly, URO (Figure 25) does not appear to be affected by the same problems as ENT and T&O, suggesting further work into understanding the differences between URO and other specialties may improve model representation.

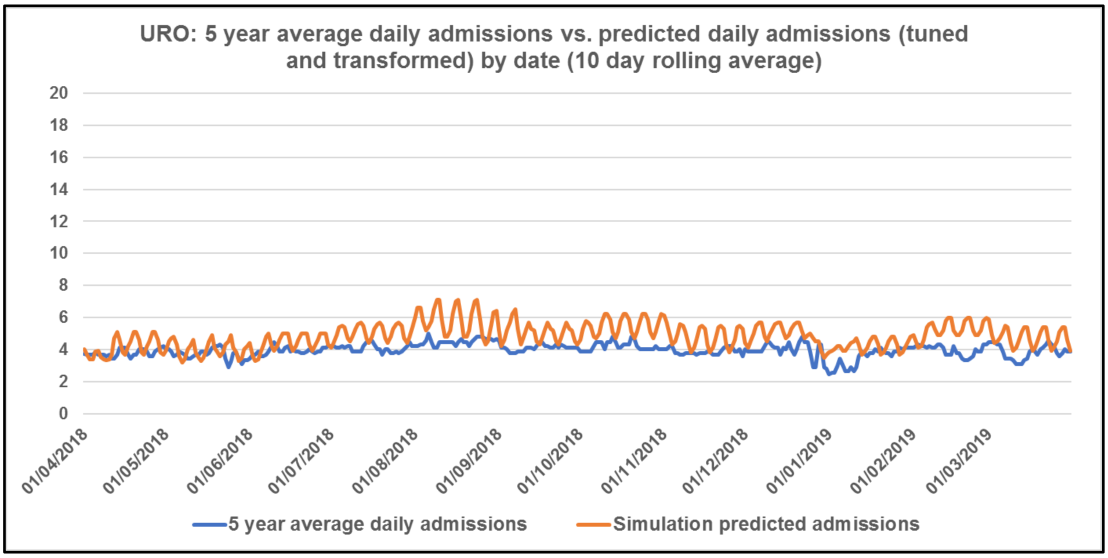

Figure 25: Tuned URO daily admissions

5.3. Analysis of results (test III)

The results in Table 17-Table 19 demonstrate that the BPDS (with a typical reduction to lost patient hours of 73-85%) performs better than all other discharge solutions developed as part of component 4. GADS1 and GADS2 (with typical reductions of 61-69%) both perform well compared to DMDS, and relatively well compared to BPDS.

Table 17: Patient lost hours (ENT)

	DMDS	GADS1	GADS2	BPDS
Lost patient hours (median)	**21,501**	**6,873**	**7,148**	**3,609**
Percentage decrease to lost patient hours	-	68.0%	66.7%	83.2%

Table 18: Patient lost hours (T&O)

	DMDS	GADS1	GADS2	BPDS
Lost patient hours (median)	**28,760**	**8,916**	**9,188**	**4,156**
Percentage decrease to lost patient hours	-	69.0%	68.1%	85.5%

Table 19: Patient lost hours (URO)

	DMDS	GADS1	GADS2	BPDS
Lost patient hours (median)	**8,366**	**2,833**	**3,194**	**2,259**
Percentage decrease to lost patient hours	-	66.1%	61.8%	73.0%

The results demonstrate that of all the solutions evaluated BPDS remains the optimal discharge solution, which is not unexpected. The results also provide sufficient merit that GADS solutions (using service desired constraints) may offer a viable alternative approach to setting fixed discharge targets, particularly since each specialty utilises unique processes with differing requirements. It is worth considering that more complicated constraints may not have such a positive impact upon lost patients hours since GADS1 and GADS2 both rely on skew yet are heavily drawn towards constraint C3 since the penalty for sum of patient lost hours is often larger than other constraints with exception to C1. This means that a discharge-profile for GADS will often be much closer in distribution to BPDS than DMDS, hence the positive results.

With the caveat that more suitable constraints should be considered, the application of test III using constraints C1-C4.2 for GADS1 and GADS2, and best practice methods (BPDS) finds that component 4 of this framework, used in conjunction with components 1, 2 and 3 is effective at reducing lost patient hours.

5.4. Analysis of results (test IV)

The results in Table 20-Table 22 demonstrate that the BPDS (with typical increases to the total number of patients discharged between 3.6%-12.6%) performs more efficiently than all other discharge solutions. GADS1 and GADS2 (with typical increases of 0.4%-12.5%) both perform more efficiently than DMDS, and relatively well compared to BPDS.

Table 20: Total patients discharged (ENT)

Measure	DMDS	GADS1	GADS2	BPDS
Total patients discharged (median)	**1,942**	**2,185**	**2,116**	**2,186**
Percentage increase of patients discharged	-	12.5%	9.0%	12.6%

Table 21: Total patients discharged (T&O)

	DMDS	GADS1	GADS2	BPDS
Total patients discharged (median	**2,318**	**2,419**	**2,328**	**2,402**
Percentage increase of patients discharged	-	4.4%	0.4%	3.6%

Table 22: Total patients discharged (URO)

	DMDS	GADS1	GADS2	BPDS
Total patients discharged (median	**1,356**	**1,395**	**1,415**	**1,415**
Percentage increase of patients discharged	-	2.9%	4.3%	4.3%

The insight offered by these results is much the same as with test III, particularly referencing the concerns about a discharge-profile bias due to constraint C3. The results of test IV demonstrates another risk of GADS which is operational consistency. Since GADS evolves a good (yet not always optimal) solution, there is no certainty that consistency will occur. Figure 26 demonstrates that while the inter-quartile range for BPDS is compact (suggesting some stability), the whiskers for both GADS are elongated in one direction more than the other which suggests increased variability.

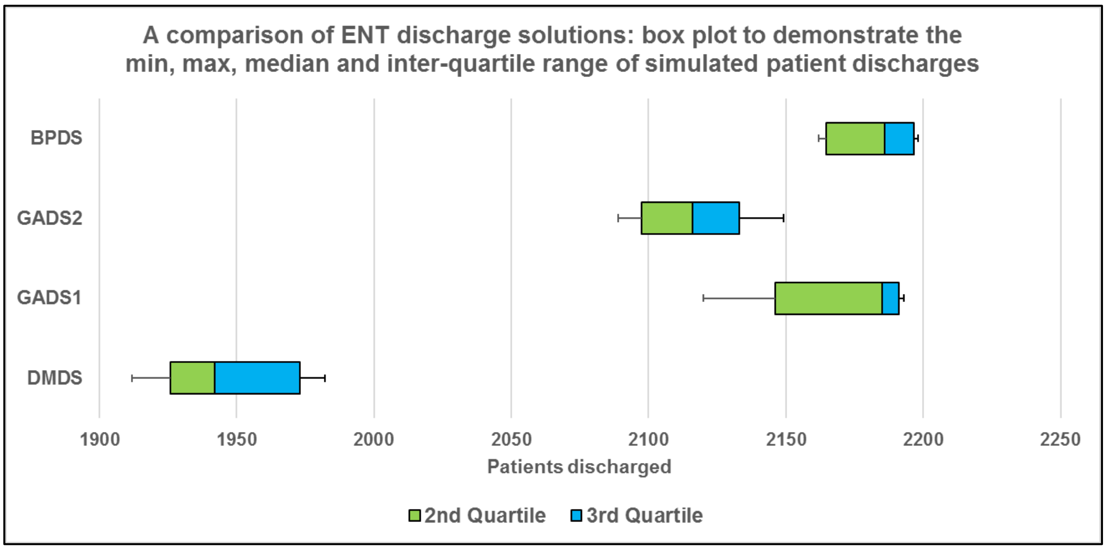

Figure 26: ENT simulated patient discharges

GADS2 in Figure 27 has an inter-quartile range which demonstrates that GADS2 can perform as effectively as BPDS and GADS1, or as ineffectively as DMDS. GADS1 has a higher median value than BPDS, however the maximum values for BPDS are higher than all other solutions. Figure 28 which displays URO results offers more consistency between discharge solutions, although with an interesting observation that while BPDS retains the best maximum number of discharges, the median of GADS2 is higher than BPDS and the range of results is more consistent. A further observation is that GADS2 performed better than GADS1 which is unexpected since the skew of GADS2 should provide more later-day discharging, thus reduce patient flow. It is possible that this is caused by repeated constraint violation, although the reasons are unknown.

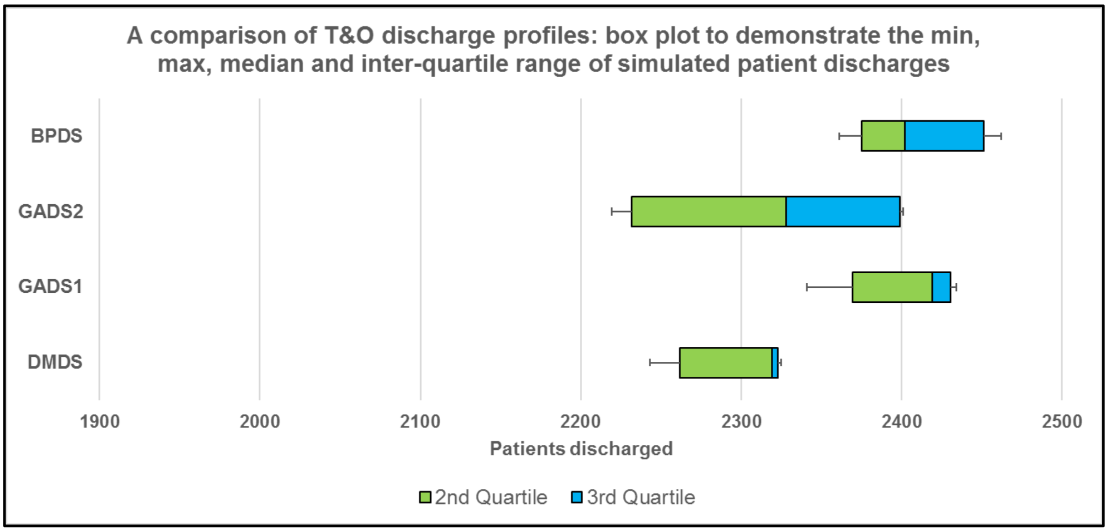

Figure 27: T&O simulated patient discharges

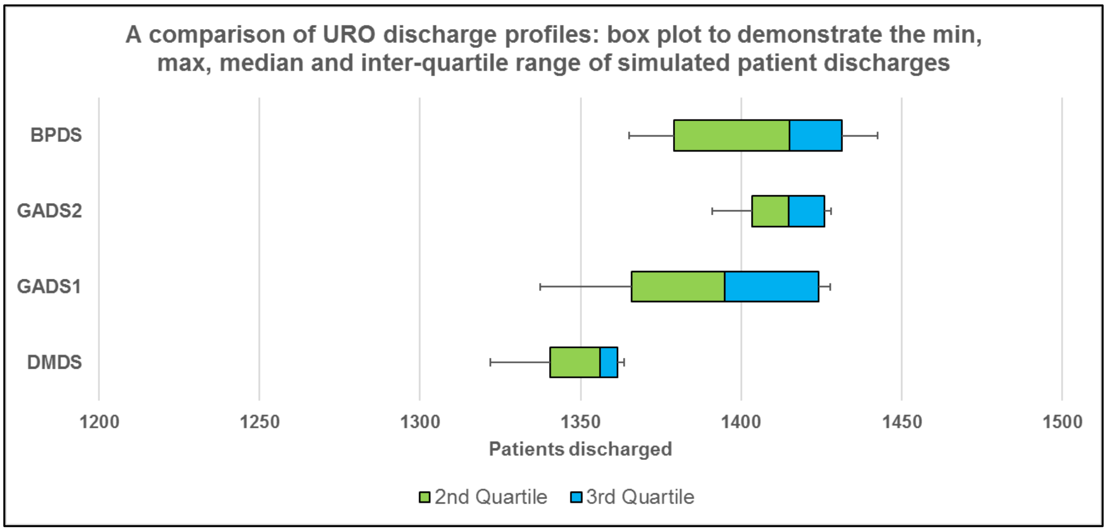

Figure 28: URO simulated patient discharges

5.5. Attainment of objectives and requirements

The development of component 1 satisfies sub-objective 3.1, while component 2 satisfies sub-objective 3.2. Component 3 satisfies sub-objective 3.4 and component 4 satisfies objective 3.3.

It is the opinion of the author that the results of test I (obtained by applying components 1 and 2 of the framework) form a solution which partially satisfies question 1 (thus objectives 3.1 and 3.2). The models provide accurate predictions against test data, with improvements possible that are within the scope of this framework. The real-world application of these models (in relation to simulation) requires the application of test II for question 1 to be satisfied. Test II found that the models provide sufficient representation of historical data with the caveat that los-model performed least-well in test I and was untested during test II. There is also further work needed to improve the accuracy of some models including further understanding of internal processes if models are to become more representative. With these caveats, the author believes that question 1 is satisfied and that objectives 3.1 and 3.2 are met.

Component 3 provides effective foundations for the framework to simulate patient flow through a specialty, with added usefulness that the system developed for component 3 can be adapted for various levels of granularity. Component 3 builds upon components 1 and 2 to meet objective 3.4. The introduction of component 4, and the analysis undertaken in test III and test IV provides sufficient evidence that the requirements set in objective 4 is achieved since discharge solutions are evaluated and BPDS and GADS are found to be effective.

The development of GA satisfied objective 3.3, thereby completing all of objective 3, however it could be argued that C4.1 and C4.2 are not suitable operational constraints since skew was utilised to provide balance. Even so, GADS1 and GADS2 performed relatively well compared to BPDS, and much better than DMDS whilst maintaining some balance to the discharge-profile thus the author considers objective 3.3.1. complete with the caveat that improvements to constraints can be attained. Objective 3.3.2. is certainly achieved since GA provides the ability for simple (or flexibility for complex) constraints to be integrated with the operational, service and patients’ needs for the purpose of simulation.

Since objectives 1-4 map directly to OR 1-6, all OR requirements are considered complete.

5.6. Limitations

The most significant limitation to this project is the constraints provided to the GA, particularly C2-C4.2 which uses skew and maximum hourly slots to balance the discharging requirements of the service against the number of lost patient hours which meets the needs of patients. The use of feasible patients, max hourly discharges, lost patient hours and distribution curve skew offers a simple representation of a complex problem, thus improved constraints based upon the unique requirements of each service would be more operationally effective and appropriate. While these constraints limit the results, the purpose of this project is to develop a framework which could be expanded on by further work.

Data quality is another known limitation, where there is no certainty that the data recorded is valid. A known issue for HDdULHB is that discharge datetimes in the patient administration system defaults to the datetime of when the record is created and are not always validated by the data entry clerks, meaning that the time (or even date) of a patients discharge may be invalid thus limiting the accuracy of the model. This limitation is reduced by restricting the number of data attributes required for developing the model, and by selecting only GGH since a recent drive to improve data quality revealed a high-degree of valid admission and discharge datetimes.

This project is constrained through using a black-box method to represent patient pathways since no consideration is given to any existing internal processes. All processes are simply represented though a patient LOS which makes it unlikely to truly reflect the service. The project excludes transfers between specialties, does not consider specialties with multiple wards or differing elective and emergency pathways, with the latter potentially affecting all three of the selected specialties. It is worth noting that the framework does not restrict users to this method and is designed as a foundation to be built upon, adding complexity and detail where necessary.

While more data is generally viewed as a positive outcome, the developed models may offer a limited representation of real-world systems since healthcare services reconfigure from time-to-time. HDdULHB has reconfigured some services during the five-year data period meaning that data from 2013 may not be as representative as data from 2018. This limitation was reduced by selecting three specialties which received little reconfiguration, however service reconfiguration and transformation work (with the latter being temporary initiative to improve an aspect of care) needs to be considered for anyone following this framework.

Another limitation is that unlike GADS and BPDS, DMDS does not contain a feature to represent the number of patients waiting for discharge. Whether this causes too few or too many discharges when performing simulations using discharge-model is unknown.

A final limitation is the computational power required to develop the models and simulate a hospital setting. The researcher used an Intel Core i5-6500 CPU @ 3.20GHz, utilising all 4-cores simultaneously, with 16GB RAM and Ubuntu 18.04.1 LTS. Each model required up to 30-hours to recreate for each cycle of hyper-parameter tuning. It is possible that with sufficient computational power, more appropriate hyper-parameters could be explored with additional accuracy gains. Alternatively, a restriction of data could be utilised in order to use larger ANNs.

Chapter 6. Conclusion and further work

This chapter brings this project to a conclusion by highlighting key themes and key points discovered as part of the project process. This chapter begins with the main body of the conclusion, before discussing recommendations for further work and finally discussing contributions to knowledge.

6.1. Conclusion

This projects aim is to explore machine learning methods and technologies to improve patient flow through a hospital setting while considering operational, staffing and patient needs. The aim is achieved with the introduction of a machine learning framework to model and optimise patient flow. The framework is applied to three specialties at HDDULHB, namely ENT, T&O and URO, and evaluated against a series of tests. The test results demonstrate that this framework can be applied to a healthcare setting and produce accurate and representative models which can be used to optimise patient flow when used in conjunction with service-level objectives. These components can be further used to simulate and evaluate the impact of several discharge solutions.

The framework is a logical conclusion of the problem domain, where the deduction from the main body is that healthcare providers need to do more activity to cope with increasing demands and will need to accomplish this prospect with reduced resource, and greater expectations from patients, internal organisations and external organisations. One method to realise this prospect is to facilitate better demand and capacity planning, particularly in relation to discharge planning which is deemed integral to optimising patient flow. There is also a growing appetite for more effective planning from both services and policy makers, with the latter starting to mandate improved planning by implementing strategic frameworks. The use of machine learning to solve this problem is promising, particularly due to relative success within other fields (which also need to facilitate good demand and capacity planning) although the use of such methods in healthcare is limited. Whilst the use of ML modelling and optimisation in healthcare is limited, the need for better service-level and process-level analysis for organisations such as HDdULHB (particularly following restructuring in 2009) has never been greater and if successfully adopted and implemented, may offer robust solutions to alleviate some of the growing pressures faced by the organisation.

This project explores the concepts of capacity and demand modelling in healthcare, particularly in relation to patient flow with the aim of optimising patient flow through HDdULHB hospitals while also providing methods which can be generalised to other hospital settings. This project also explores the use of machine learning and simulation to represent the model of flow, and to evaluate discharge solutions. This project provides a machine learning framework for developing and optimising models for the purpose of simulating patient flow through a hospital setting. The framework is evaluated for effectiveness of model development using two tests derived from the following question:

"Is the framework effective for developing models of patient flow through a hospital setting?"

To which the author believes – based on the resulting data – is yes. The two tests found that most models predicted admission, LOS and discharges against test data with greater than 75% accuracy, and up to 90% accuracy. The tests also found the data to be representative of historical profiles, however further work is required to improve accuracy and representation, particularly for los-model and for predicting seasonal variation. This can be achieved by improving the methods which form components 1 and 2 of this framework.

This project provides several discharge-profiles to control the flow out of a hospital setting. The framework includes the ability to evaluate discharge-profiles (component 4) for effectiveness, thus optimising patient flow by using simulation (component 3). The framework is evaluated for effectiveness of optimising patient flow using two tests derived from the question:

"Is the framework effective for optimising patient flow *through a hospital setting?"*

To which the author also believes is yes following analysis of the resulting data.

It is evident from the test results that the discharge solution of 80% of patients before 11am and 100% before 12pm (BPDS) is best practice for a reason. The results demonstrate – in conjunction with research discussed in the literature – that improved patient flow of 3.6-12.6% is achieved by adopting to this method. The reality found in the literature, and based upon the authors experience, is that this target will continue to be difficult to achieve, with calls from clinicians to provide a more achievable discharging methods.

The test results demonstrate that a GADS – with increased total discharging, thus flow of between 0.4% and 12.5% – provides a viable alternative method for discharging patients when compared to BPDS and DMDS since GADS provides the ability to balance conflicting demands and prioritise the objectives of each specialty service. This facilitates some flexibility and (if used correctly with suitable constraints) could facilitate safer discharging, both of which are requirements desired by the service, although a key risk of deploying GADS is that operational consistency may not be achieved, since GA produce good, though not always best solutions. This is evident is some results where the inter-quartile range of results of GADS2 ranged from being the worst solution to be the best solution.

Whilst the simple GA objectives used as part of this project provide some success, it is imperative that users of this framework explore the real objectives faced by each elements of the pathway which they wish to optimise. Realistically, skew is not sufficient to balance a discharge-profile but serves as a proof of concept which solution could be viewed as a stepping stone to achieving the best practice solution.

Objectives 1 and 2 of this project are achieved by identifying discharge-profile as a metric and ML as a technology to improve patient flow through a hospital setting. Objective 3 is achieved by developing components 1-4 of the framework and applying the components to tests I and II. Objective 4 is achieved by applying components 1-4 to test III and IV. This means that OR1-6 are achieved on behalf of HDdULHB.

6.2. Recommendations for further work

Throughout this project, several limitations are identified which provide the basis for further work. The use of a ML framework to model and optimise patient flow is promising, however healthcare providers are typically complex organisations with a large number of internal processes which are not well documented, meaning that simple models of flow will not truly reflect the service. In relation to further developing systems to be used with the framework, there are three themes of further work:

1. Improvements to the objective functions and constraints which provide better (or more applicable) discharge-profiles based upon the specific requirements of each service.
2. Improvements to the simulation environment with more detailed internal processes.
3. Improvements to model representation through attribute identification or tuning.

With respect to point 1, the use of skew alone is insufficient to realistically balance a discharge-profile which may be better served by mapping constraints to patient categories (i.e. Pharmacy requirements) which are prioritised, or matching discharge-profiles to ward rounds. This project utilised four key constraints to ensure that GADS discharged all available patients yet did not provide more slots than required. The constraints also ensure that the number of patient lost hours is minimised. This experiment did not consider multiple discharging stages throughout the day (which could mimic ward rounds), nor did it consider constraints which made small improvements to current discharge patterns (as identified in the literature) which could help facilitate small step changes for a service on the road to achieving best practice discharging. This project failed to address a problem mentioned as part of the problem statement in chapter 3.1 since the evaluation did not apply a balancing of discharge slots to the number of admissions. A key reason for omitting this problem is the complexity of developing this framework left insufficient time to achieve it. This framework can be built upon and facilitates the generation of objectives; thus, it is possible to use admission-model to predict demand which is then provided to a GADS objective which balances the number of discharge slots to predicted admissions and prioritises patients based upon their probability of timely discharge. This would require further work and study as to how such patients could be identified and prioritised, although there are suggestions in the literature around age related delays being prevalent.

With respect to point 2, a more detailed specialty pathway can be built using the foundations provided by this framework, particularly generating more detailed patient flows which include transfers to other wards, known delays to discharging (such as medication delays), simulated multi-disciplinary team ward rounds and segregating emergency and elective pathways. By improving the detail of the pathway simulation, it will be possible to more accurately evaluate the outcomes of potential discharge solutions, thus gain a better understanding of patient flow.

Finally, improvements relating to point 3 include generating better models which represent demand and activity. A few observations made from tests I-IV is that additional attributes are required. Some examples include the need to identify specialty processes and include them in the models. i.e., analysis suggests that elective and emergency patients are reducing model accuracy during annual temporal events. Secondly, los-model obtained low accuracy for some specialties, this suggests that further investigation into suitable LOS attributes for predicting patient LOS (such as presenting conditions) is required.

6.3. Contributions to knowledge

This project contributes further evidence to suggest that 80% of patients discharged before 11am is an optimal discharge solution and remains best practice with increased flows through a simulated specialty of up to 12.6% observed. This project also contributes to knowledge that GA can be applied in conjunction with simulation and ML developed capacity and demand models to optimise patient flow, subject to specialty level constraints. Such methods applied as part of this project observed increased simulated flow through a specialty of up to 12.5%.

This project has contributed to knowledge that SG methods can be applied to effectively model admissions, LOS and discharges, building upon demand modelling work undertaken by Tugay & Oguducu (2017) with accuracy scoring against test data of up to 90%, although further work is required to realise LOS models. In most instances, SG is in the top two classifiers for accuracy, and analysis of admission-profiles found SG to provide a high-degree of hourly and daily prediction representation. Finally, this project has contributed to knowledge by bringing together existing components of database, data mining and machine learning and forming a framework which can be extrapolated to many healthcare settings.

Chapter 7. Reflection

This chapter provides my personal reflection of the journey through this project and dissertation. I started this project with the broad aim of increasing patient flow through a hospital setting, initially believing that this would be achieved through either predicting emergency hospital demand temporally (thus facilitate planning of staffing levels) by combining machine learning techniques with HDdULHB data, local event data and weather data. Another idea was to use machine learning techniques to facilitate the analysis of discharge patterns and identify cases of good (and bad) practice, which could later be used by hospital services to enhance local processes. I was given a broad organisation aim of identifying new performance metrics or establishing some method to facilitate better hospital planning or patient flow. In hindsight, I would have narrowed down the aims scope, and agreed it with HDdULHB at an earlier stage to limit the amount of literature I reviewed, much of which turned out not to be relevant to this project. For future projects, I will ensure that a clear (and agreed) aim is established between myself and my sponsor at an early stage.

As I delved further into the literature, it became clear that there was a need to improve discharging, particularly shifting the discharge-profile distribution curve forward so that more patients are discharged earlier in the day, theoretically increasing patient flow through an entire hospital and reducing ED crowding. Even though evidence suggests that discharging prior to midday offered the most effective results, with my personal experience working for several NHS organisations I felt that this would be difficult to achieve, and I empathised with calls in the literature from clinicians, whom desired flexibility. I decided that this would be an interesting problem to solve, and quickly realised that the problem could be best solved using genetic algorithms to try to balance discharging and demand based on several constraints. I feel that I identified a suitable problem, and potential solutions quite early on so my methods of reviewing literature will continue into future projects.

One of the most difficult aspects of this project was mapping the required data warehouse attributes to the desired dataset using a fit-for-purpose methodology. While there is sufficient suggestions in the literature as to what the attributes should be, there are several complexities to building a dataset which were not mentioned in the literature. i.e. it took several iterations of development before I realised that I only needed to include inpatients and exclude patients who had passed away since both groups follow different discharge processes. In future, I will ensure that I meet with staff at the information team at an earlier stage, since their insight helped me narrow down the methodology more quickly.

An early error of judgment was the failure to fully consider evaluation methods prior to system development (despite advance warning from my supervisor), and while this did not slow down my development work, I needed to retrospectively adjust some of the methods used to something more sensible which required further research of several topics in the literature. This could have been avoided if I was aware of the evaluation methods before beginning development, or at least a high-level conceptual idea of the methods. A key reason for this error is that I was not sure what was really meant by evaluation methods, whether this was the experiment results, the evaluation of the project, objectives, all of which I felt that I had sufficient idea of what they would be, but in future, I will ensure that I note all evaluation criteria for all aspects at an early stage so that I always have the end-goal in mind.

I also failed to comprehensively develop my objectives (partly due to the broad organisation requirements, and partly due to me not narrowing the scope) and map them to the components of the framework at an early stage. This meant that aspects of my report must have appeared disjointed to my supervisor since it felt that way to me and resulted in a need to invest a lot of time changing elements of the report, backtracking to join up the flow of the dissertation which could have been avoided if I was more methodical at the beginning. Time management was difficult throughout the year due to personal reasons requiring me to balance family time, work time, project time and even home renovation time. I overcame this challenge by ensuring that there was ringfenced time for me to work on this project at key moments of the year. In future, I will create a more realistic schedule and agree suitable “project slots” with my family and colleagues.

One element I was particularly pleased with was how the entire syllabus of my MSc programme somehow came together in this project. I studied concepts such as distance measures, GA, hyper-tuning, functional programming, categorisation to name a few without fully-comprehending where these concepts can be applied.

During this project, I felt many “eureka” moments where I knew what the right tool for the job would be, such as applying distance measures as GA penalty constraints.

A final few related points for reflection is that I felt that I was bringing together several important components to solve a problem identified in the literature, but even though I felt that I knew what I was doing, it was difficult to write about because it did not have a name. I wondered if it was a system? An application? A method? It was only when I realised that I was creating a framework that everything made sense and I was able to make the report flow well. I thank my supervisor for pointing out the need to name objects to make sense of them, and also applied this idea to mathematical symbols, framework concepts such as components or specific models such as los-models. This experience has challenged me mentally but has also been very interesting and I feel that I have accomplished and learnt a lot by completing this project, this experience will help with future report writing.

References

Allen, I., 2001. Stress Among Ward Sisters and Charge Nurses, London: Policy Studies Institute.

Aridas, C. & Lemaitre , G., 2016. Pipeline Object. [Online]
Available at: http://contrib.scikit-learn.org/imbalanced-learn/stable/auto_examples/pipeline/plot_pipeline_classification.html#sphx-glr-auto-examples-pipeline-plot-pipeline-classification-py
[Accessed 18 Aug 2018].

Blagus, R. & Lusa, L., 2013. SMOTE for high-dimensional class-imbalanced data. BMC Bioinformatics, 14(2013), p. 106.

Bourn, J. et al., 2003. Ensuring the effective discharge of older patients from NHS acute hospitals, London: National Audit Office.

British Medical Association, 2016. State of the health system - Beds in the NHS: Wales, London: British Medical Association.

Bruni, R. & Detti, P., 2014. A flexible discrete optimization approach to the physician scheduling problem. Operations Research for Health Care, 3(2014), pp. 191-199.

Cao, X. H., Stojkovic, I. & Obradovic, Z., 2016. A robust data scaling algorithm to improve classification accuracies in biomedical data. BMC Bioinformatics, 17(1), p. 359.

Care Quality Commission, 2017. The state of health care and adult social care in England 2016/17, London: Care Quality Commission.

Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P., 2002. SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research, 16(2002), pp. 321-357.

Chun-Chin, W., Tian-Shy, L. & Kuo-Liang, L., 2008. An ERP performance measurement framework using a fuzzy integral approach. Journal of Manufacturing Technology Management, 19(5), pp. 607-626.

Claesen, M., Simm, J., Popovic, D. & De Moor, B., 2017. Hyperparameter tuning in Python using Optunity, Leuven: KU Leuven.

Connolly, M. et al., 2009. Systems and people under pressure: the discharge process in an acute hospital. Journal of Clinical Nursing, 18(4), pp. 549-558.

DEAP Project, 2018. DEAP documentation. [Online]
Available at: https://deap.readthedocs.io/en/master/
[Accessed 1 May 2018].

DEAP, 2018. Constraint Handling. [Online]
Available at:
https://deap.readthedocs.io/en/master/tutorials/advanced/constraints.html
[Accessed 8 June 2018].

Dittman, D. J., Khoshgoftaar, T. M., Wald, R. & Napolitano, A., 2014. Comparison of Data Sampling Approaches for Imbalanced Bioinformatics Data. Pensacola Beach, Florida Artificial Intelligence Research Society Conference.

Evans, N., 2017. Is the NHS heading for the worst winter in a generation?. Nursing Management, 24(6), p. 8.

Health in Wales, 2009. Structure. [Online]
Available at: http://www.wales.nhs.uk/nhswalesaboutus/structure
[Accessed 4 December 2017].

Healthcare Commision, 2007. State of Healthcare 2007: Improvements and challenges in services in England and Wales, London: Commission for Healthcare Audit and Inspection.

Healthcare Commission, 2008. Spotlight on Complaints, London: Commission for Healthcare Audit and Inspection .

Healthcare Inspectorate Wales, 2015. Hospital Inspection (Unannounced): Hywel Dda University Health Board, Merthyr Tydfil: Healthcare Inspectorate Wales.

Hertzum, M., 2016. Patterns in Emergency-Department Arrivals and Length of Stay: Input for Visualizations of Crowding. The Ergonomics Open Journal, 9(2016), pp. 1-11.

Jones, S. S. et al., 2009. A multivariate time series approach to modeling and forecasting demand in the emergency department. Journal of Biomedical Informatics, 42(1), pp. 123-139.

Khanna, S., Sier, D. & Zeitz, K., 2016. Discharge timeliness and its impact on hospital crowding and emergency department flow performance. Emergency Medicine Australasia, 28(2), pp. 164-170.

Lemaitre, G., Nogueira, F., Oliveira, D. & Aridas, C., 2016. imbalanced-learn API. [Online]
Available at: http://contrib.scikit-learn.org/imbalanced-learn/stable/api.html
[Accessed 30 July 2018].

Lin, C. I., 2008. Optimization Models for Capacity Planning in Health Care, Gainesville: University of Florida.

Liu, S. W. et al., 2009. A Pilot Study Examining Undesirable Events Among Emergency Department–Boarded Patients Awaiting Inpatient Beds. Annals of Emergency Medicine, 54(3), pp. 381-385.

Maramba, P. J., Richards, S., Myers, A. L. & Larrabee, J. H., 2004. Discharge Planning Process: Applying a Model for Evidence-based Practice. Journal of Nursing Care Quality, 19(2), pp. 123-129.

Mishra, S., 2017. Handling Imbalanced Data: SMOTE vs. Random Undersampling. International Research Journal of Engineering and Technology, 4(8), pp. 317-320.

Mutingi, M. & Mbohwa, C., 2013. A Fuzzy Genetic Algorithm for Healthcare Staff Scheduling. Johannesburg, International Conference on Law, Entrepreneurship and Industrial Engineering.

NHS Direct Wales, 2017. Glangwili General Hospital. [Online]
Available at:
https://www.nhsdirect.wales.nhs.uk/LocalServices/ViewLocalService.aspx?id=3185&s=Hospital
[Accessed 10 November 2017].

Powell, E. S. et al., 2012. The relationship between inpatient discharge timing and emergency department boarding. Journal of Emergency Medicine, 42(2), pp. 186-196.

Puente, J., Gómez, A., Fernández, I. & Priore, P., 2009. Medical doctor rostering problem in a hospital emergency department by means of genetic algorithms. Computers & Industrial Engineering, 56(2009), pp. 1232-1242.

Raghav, R. V., Guillaume, L. & Thomas, U., 2017. 4.3. Preprocessing data. [Online]
Available at: http://scikit-learn.org/stable/auto_examples/preprocessing/plot_all_scaling.html
[Accessed 8 Apr 2018].

Raschka, S., 2018. StackingCVClassifier. [Online]
Available at:
https://rasbt.github.io/mlxtend/user_guide/classifier/StackingCVClassifier/
[Accessed 29 May 2018].

Rodrigues , F., Cardeira, C. & Calado, J. M. F., 2014. The daily and hourly energy consumption and load forecasting using artificial neural network method: a case study using a set of 93 households in Portugal. Energy Procedia, 62(2014), pp. 220-229.

Ryu, S., Noh, J. & Kim, H., 2016. Deep Neural Network Based Demand Side Short Term Load Forecasting. Energies, 10(1).

scikit-learn developers, 2017. Classifier comparison. [Online]
Available at: http://scikit-learn.org/stable/auto_examples/classification/plot_classifier_comparison.html
[Accessed 19 Apr 2018].

scikit-learn developers, 2017. Model evaluation: quantifying the quality of predictions. [Online]
Available at: http://scikit-learn.org/stable/modules/model_evaluation.html
[Accessed 10 Apr 2018].

scikit-learn, 2017. sklearn.pipeline.Pipeline. [Online]
Available at: http://scikit-learn.org/stable/modules/generated/sklearn.pipeline.Pipeline.html
[Accessed 12 Jun 2018].

Smith, A. et al., 2014. Predictors of Emergency Department breaches in patients admitted onto a Medical Admissions Unit. Manchester, The University of Manchester.

Sørup, M. C., Jacobsen, P. & Forberg, J. L., 2013. Evaluation of emergency department performance – a systematic review on recommended performance and quality-in-care measures. 21(62).

Statistics for Wales, 2011. Health Trends in Wales, Cardiff: Welsh Assembly Government.

Stolletz, R. & Brunner, J. O., 2012. Fair optimization of fortnightly physician schedules with flexible shifts. European Journal of Operational Research, 2019(2012), pp. 622-629.

Subramanian, N., 2013. A Strategic Capacity Planning Model: A Genetic Algorithm Approach. International Journal Of Engineering And Computer Science, 2(12), pp. 3455-3464.

Szoplik, J., 2015. Forecasting of natural gas consumption with artificial neural networks. Energy, 85(2015), pp. 208-220.

Team SimPy, 2017. Shared resource primitives. [Online]
Available at:
https://simpy.readthedocs.io/en/latest/api_reference/simpy.resources.html
[Accessed 2018 Jun 13].

Team SimPy, 2018. Overview. [Online]
Available at: https://simpy.readthedocs.io/en/latest/
[Accessed 6 Apr 2018].

Thomas, S. H., 2017. *Redesigning 'step up' and hospital 'front door' services to meet* the needs of older patients, Cardiff: School of Mathematics, Cardiff University.

Tugay, R. & Oguducu, S. G., 2017. Demand Prediction Using Machine Learning Methods and Stacked Generalization. Madrid, DATA'17.

Wargon, M., Guidet, B., Hoang, T. D. & Hejblum, G., 2009. A systematic review of models for forecasting the number of emergency department visits. Emergency Medicine Journal, 26(3), pp. 395-399.

Welsh Assembly Government, 2009. NHS Wales: Why we are changing the structure, Cardiff: Welsh Assembly Government.

Welsh Governemnt, 2016. NHS Wales Outcomes Framework 2016-2017. Cardiff: OGL.

Welsh Government, 2017. NHS activity and performance summary. [Online]
Available at: http://gov.wales/statistics-and-research/nhs-activity-performance-summary/?skip=1&lang=en
[Accessed 30 November 2017].

Welsh Government, 2017. NHS Wales Planning Framework 2018/21, Cardiff: Welsh Government.

Welsh Government, 2017. Performance against 12 hour waiting times target by hospital. [Online]
Available at: https://statswales.gov.wales/Catalogue/Health-and-Social-Care/NHS-Hospital-Waiting-Times/Accident-and-Emergency/performanceagainst12hourwaitingtimestarget-by-hospital
[Accessed 1 November 2017].

Welsh Government, 2017. Performance against 4 hour waiting times target by hospital. [Online]
Available at: https://statswales.gov.wales/Catalogue/Health-and-Social-Care/NHS-Hospital-Waiting-Times/Accident-and-Emergency/performanceagainst4hourwaitingtimestarget-by-hospital
[Accessed 2 November 2017].

Witten, I. H., Frank, E., Hall, M. A. & Pal, C. J., 2017. Data Mining: Practical Machine Learning Tools and Techniques. 4th ed. London: Morgan Kaufmann.

Wolpert, D. H., 1992. Stacked Generalization. Neural Networks, 5(1992), pp. 241-259.

Xu, M., Wong, T. C. & Chin, K. S., 2013. Modeling daily patient arrivals at Emergency Department and quantifying the relative importance of contributing variables using artificial neural network. Decision Support Systems, 54(2013), p. 1488–1498.

Yeh, J. & Lin, W., 2007. Using simulation technique and genetic algorithm to improve the quality care of a hospital emergency department. Expert Systems with Applications, 32(4), pp. 1073-1083.

Appendix I - Table of abbreviations

Abbreviation	Full name
ALOS	Average Length of Stay
ANN	Artificial Neural Network
ASDS	Admitted Spells Dataset
BPDS	Best Practice Discharge Solution
CXPB	Crossover Probability
DEAP	Distributes Evolutionary Algorithms in Python
DM	Data Mining
DMDS	Discharge-model Discharge Solution
DT	Decision Tree
DTOC	Delayed Transfer of Care
ED	Emergency Department
ENT	Ear, Nose and Throat
FIFO	First In First Out
GA	Genetic Algorithm
GADS	Genetic Algorithm Discharge Solution
GGH	Glangwili General Hospital
GB	Gradient Boosting
HB	Health Board
HDDULHB	Hywel Dda University Local Health Board
LHB	Local Health Board
LOS	Length of stay
LR	Logistic Regression
ML	Machine Learning
MLP	Multi-Layer Perceptron
MUTPB	Mutation Probability
NHS	National Health Service
OR	Organisational Requirement
RF	Random Forest
SG	Stacked Generalisation
SMOTE	Synthetic Minority Oversampling Technique

T&O	Trauma & Orthopaedic
URO	Urology
WG	Welsh Government

Appendix II – Data attributes for admission, LOS and discharge datasets

Field	Data Type	Dataset
Hour0	BIT (1 or 0)	All
Hour1	BIT (1 or 0)	All
Hour2	BIT (1 or 0)	All
Hour3	BIT (1 or 0)	All
Hour4	BIT (1 or 0)	All
Hour5	BIT (1 or 0)	All
Hour6	BIT (1 or 0)	All
Hour7	BIT (1 or 0)	All
Hour8	BIT (1 or 0)	All
Hour9	BIT (1 or 0)	All
Hour10	BIT (1 or 0)	All
Hour11	BIT (1 or 0)	All
Hour12	BIT (1 or 0)	All
Hour13	BIT (1 or 0)	All
Hour14	BIT (1 or 0)	All
Hour15	BIT (1 or 0)	All
Hour16	BIT (1 or 0)	All
Hour17	BIT (1 or 0)	All
Hour18	BIT (1 or 0)	All
Hour19	BIT (1 or 0)	All
Hour20	BIT (1 or 0)	All
Hour21	BIT (1 or 0)	All
Hour22	BIT (1 or 0)	All
Hour23	BIT (1 or 0)	All
Jan	BIT (1 or 0)	All
Feb	BIT (1 or 0)	All
Mar	BIT (1 or 0)	All
Apr	BIT (1 or 0)	All
May	BIT (1 or 0)	All
Jun	BIT (1 or 0)	All
Jul	BIT (1 or 0)	All
Aug	BIT (1 or 0)	All

Sep	BIT (1 or 0)	All
Oct	BIT (1 or 0)	All
Nov	BIT (1 or 0)	All
Dec	BIT (1 or 0)	All
Mon	BIT (1 or 0)	All
Tue	BIT (1 or 0)	All
Wed	BIT (1 or 0)	All
Thu	BIT (1 or 0)	All
Fri	BIT (1 or 0)	All
Sat	BIT (1 or 0)	All
Sun	BIT (1 or 0)	All
IsBankHoliday	BIT (1 or 0)	All
IsDayAfterBankHoliday	BIT (1 or 0)	All
Winter	BIT (1 or 0)	All
Spring	BIT (1 or 0)	All
Summer	BIT (1 or 0)	All
Autumn	BIT (1 or 0)	All
NumberOfAdmissionsCatUDF	INT	Admission
NumberOfDischargesCatUDF	INT	Discharge
Male	BIT (1 or 0)	LOS
Female	BIT (1 or 0)	LOS
0to17	BIT (1 or 0)	LOS
18to64	BIT (1 or 0)	LOS
65to74	BIT (1 or 0)	LOS
75to84	BIT (1 or 0)	LOS
Over85	BIT (1 or 0)	LOS
Age	INT	LOS
LOSGroup	INT	LOS

Appendix III – List of complimentary files

File	Description
Attributes.py	Python script to control parameters for classifier, generic algorithm and pathway simulation classes.
Classifier.py	Python script to generate machine learning models for admissions, discharges and LOS. Facilitates hyper-parameter tuning and cross validation. Includes methods to make unseen data predictions, loading and saving models.
GeneticAlgorithm.py	Python script to generate a GA discharge-profile based upon defined objectives.
PathwaySimulation.py	Python script to run a pathway simulation, using admission-model and los-model, and calling upon a discharge solution to generate a discharge-profile.
RebuildModel.py	Python script to control the rebuilding of machine learning models. This class calls on the Classifier.py class.
Admission-Dataset.sql	SQL script to generate admission-dataset
Discharge-Dataset.sql	SQL script to generate discharge-dataset
LOS-Dataset.sql	SQL script to generate los-dataset

201108

NOTICE OF SUBMISSION: POSTGRADUATE TAUGHT DEGREES

CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

Please TYPE or write in BLACK ink and use BLOCK capitals

SECTION A: TO BE COMPLETED BY THE CANDIDATE AND SUBMITTED WITH THE DISSERTATION

Student ID Number:	C1573025
Title:	**Please circle appropriate value** **Mr/ Miss/ Ms/ Mrs/ Dr/ Rev/ Other, please specify ....**
Surname:	NEALE
First Names:	SEBASTIAN GWILYM
School:	COMPUTER SCIENCE & INFORMATICS
Title of Degree:	**Please circle appropriate value** **MA / MSc/ MScEcon/ MBA / MEd / LLM / MMus/ MTh / MPH / MScD / Other, please specify ....**
Full Title of Dissertation submitted:	A MACHINE LEARNING FRAMEWORK FOR MODELLING AND OPTIMISING PATIENT FLOW THROUGH A HOSPITAL SETTING
Is this a resubmission?	**Yes** ☐ **No** ☒
Dissertation submitted for Examination in:	**Permanent Binding** ☐ **Temporary binding** ☐
Supervisor's Name:	DR ANGELIKA KIMMIG
Address for receipt of your Result Letter, Degree Certificate and Details of the Graduation	20A HEOL Y FELIN

201108

Ceremony:	SEVEN SISTERS NEATH PORT TALBOT **Postcode** (if applicable): SA10 9BH
Please indicate whether you wish to attend the Degree Ceremony:	**Yes** ☐ **No** ☒
Preferred Contact Telephone (with dialling code/s):	07827 889 085
Email/s:	SEBNEALE@GMAIL.COM

CANDIDATE'S SIGNATURE: .... **DATE OF SUBMISSION:** 07/09/18.

SECTION B: TO BE COMPLETED BY THE SCHOOL

Please tick the appropriate box to confirm that the following documents have been completed by the candidate and deposited in the School.

Dissertation Summary Sheet:	**Yes** ☐	**No** ☐
Statements and Declaration:	**Yes** ☐	**No** ☐
A signed statement regarding availability of the Dissertation:	**Yes** ☐	**No** ☐
Where Candidate is resubmitting, School should attach cheque (or evidence that the fee has been paid):	**Yes** ☐	**No** ☐

SCHOOL'S SIGNATURE: .... **DATE:** ..... / ..... / ......

POSITION: ....