

# Vendor Managed Inventory Problem for Blood Platelets: a Two Stage Deep Reinforcement Learning Approach

Andres F. Osorio<sup>a</sup>, Juan D. Carvajal<sup>b</sup>

<sup>a</sup>*afosorio@icesi.edu.co*

<sup>b</sup>*juan030698@hotmail.com*

---

## Abstract

This paper presents an application of Deep Reinforcement techniques

---

## 1. Introduction

## 2. Literature Review

## 3. Problem Definition

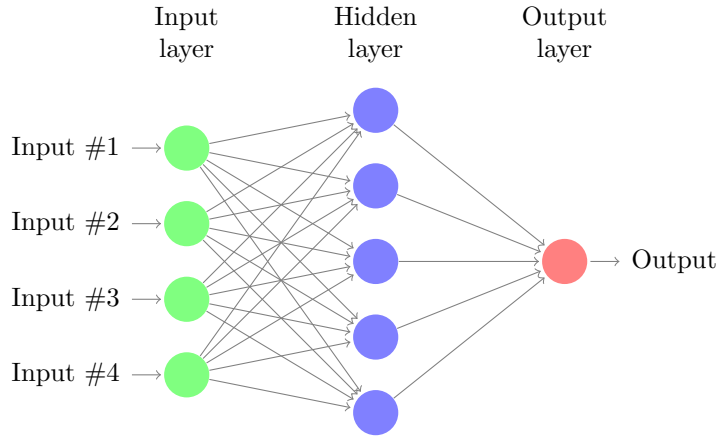
### 3.1. Context

In general terms, the supply chain of human blood products its very complicated. Blood products, specially platelets, have very short expiration periods (around 5 days). Because of this, blood managing entities , like donor centers, face great issues estimating and planning production of these vital products. These issues can be of two types, the first one, its about expiration levels, if the donor center produces too much platelets (overestimating the demand), these limited and costly to produce items will go to waste, generating high costs for the entity, and providing no benefit to final users. The second kind of issue, is related to product stock outs (underestimating the demand). Its clear that stock outs are a much more important problem for the donor centers, because these can result in increased risk of death for those needing the platelets. The gravity of this problems is only increased when multiple hospitals or medical centers, with different demand distributions are dependent on the supplies from the donor center. The basic decision for a donor center consists of two parts, the first one, its about how much to produce. Human blood can be transformed and separated into its basic constituents (platelets among these) using a process called blood fractionation, usually performed centrifuging the blood. This process is carried out using specially designed blood bags, which can be single, double, triple or quadruple. Only the last two (the most expensive ones) can be used to extract platelets, so the donor center must know how many of these bags to use in the total donor pool. The second problem its about how much to deliver to each medical center, given its demand distribution and the current inventory

of the donor center. This problem is essentially a particular case of a Vendor Managed Inventory System, where a distribution center has to supply each of its customers based on individual needs and current capacity.

### 3.2. Modeling

To model such a system, a 4 medical center distribution has been considered, but the model should be easily scaled to an arbitrary number of medical centers. The medical centers will be referred from now on as hospitals. Each hospital has an internal inventory, visible to the donor center, this inventory is represented by a 5-tuple:



## 4. Methodology

The deep reinforcement learning model was developed using the Tensorflow, the integer programming model was solved using pulp the integration was carried out using the Python

Feynman and Vernon Jr. (1963) Dirac (1953)

### 4.1. Two Stage Deep Reinforcement Learning Framework

In our framework, deep reinforcement learning is used to represent the actual of the blood centre, incorporating variability in donation and demand. Features such as perishability. Integer programming is used to support decisions concerning the allocation of blood platelets to the hospitals.

### 4.2. Deep Reinforcement Learning

$$Q(s_t, a_t) \leftarrow (1 - \alpha) \cdot Q(s_t, a_t) + \alpha \cdot (r_t + \gamma \cdot \max_a Q(s_{t+1}, a)) \quad (1)$$

#### 4.3. Integer Programming Model

##### Objective Function

$$\min \sum_{h \in H} (I_{h0}CV + F_hCF) \quad (2)$$

##### Constraints

$$I_{h0} = \max(0, I_{h1} + X_{h1} - D_h) \quad (3)$$

$$F_h = \max(0, D_h - \sum_{r \in R} (I_{hr} + X_{hr})) \quad (4)$$

$$F_h \leq \frac{1}{\|H\|} \sum_{h \in H} (F_h) \quad (5)$$

$$A_r = \sum_{h \in H} (X_{hr}) \quad (6)$$

$$\frac{\sum_{r \in R} (I_{hr} + X_{hr})}{D_h} = \frac{\sum_{r \in R} (I_{h+1r} + X_{h+1r})}{D_{h+1}} \quad \forall h < 4 \quad (7)$$

Constraints 2 and 3 can be linearised as follows:  
Linearisation Equation 2

$$-I_{h1} - X_{h1} + D_h \leq M * Y I_{h0} \quad \forall h \quad (8)$$

$$I_{h1} + X_{h1} - D_h \leq M(1 - Y I_{h0}) \quad \forall h \quad (9)$$

$$I_{h0} \geq 0 \quad \forall h \quad (10)$$

$$I_{h0} \geq I_{h1} + X_{h1} - D_h \quad \forall h \quad (11)$$

$$I_{h0} \leq M(1 - Y I_{h0}) \quad \forall h \quad (12)$$

$$I_{h0} \leq I_{h1} + X_{h1} - D_h + M * Y I_{h0} \quad \forall h \quad (13)$$

Linearisation Equation 3

$$-D_h + \sum_{r \in R} (I_{hr} + X_{hr}) \leq M * Y F_h \quad \forall h \quad (14)$$

$$D_h - \sum_{r \in R} (I_{hr} + X_{hr}) \leq M * (1 - Y F_h) \quad \forall h \quad (15)$$

$$F_h \geq 0 \quad \forall \quad h \quad (16)$$

$$F_h \geq D_h - \sum_{r \in R} (I_{hr} + X_{hr}) \quad \forall \quad h \quad (17)$$

$$F_h \leq M(1 - YF_h) \quad \forall \quad h \quad (18)$$

$$F_h \leq D_h - \sum_{r \in R} (I_{hr} + X_{hr}) + M * YF_h \quad \forall h \quad (19)$$

Once the model has been run, the inventory in each hospital is updated as follows:

$$I_{hr} = \max(0, I_{hr+1} + X_h - \max(0, D_h - \sum_{r \in R} (I_{hr} + X_{hr}))) \quad \forall h, \quad r \neq 0, 4 \quad (20)$$

$$I_{h4} = \max(0, X_4 - \max(0, D_h - \sum_{r \in R} (I_{hr} + X_{hr}))) \quad (21)$$

## 5. Case study

## 6. Results

### References

- Dirac P. The lorentz transformation and absolute time. *Physica*1953;19(1-12):888–96.
- Feynman R., Vernon Jr. F. The theory of a general quantum system interacting with a linear dissipative system. *Annals of Physics*1963;24:118–73.