

POLITECNICO DI MILANO
Corso di Laurea Magistrale in Ingegneria Biomedica
Dipartimento di Elettronica, Informazione e Bioingegneria



**LA MANUTENZIONE DELLE
APPARECCHIATURE BIOMEDICHE
NELL'ERA DELLE TECNOLOGIE DIGITALI**

Relatore: Prof. Veronica Cimolin

Correlatore: Ing. Daniela Motta

Tesi di Laurea di:
Francesco Raffaele Tranquillo, matricola 905980

Anno Accademico 2018-2019

Spazio per dedica

Indice

Sommario	V
Abstract	VI
1 Introduzione e contesto	1
1.1 Organizzazione della ASST Vimercate	1
1.1.1 Polo Ospedaliero	2
1.1.2 Rete Territoriale	2
1.1.3 La Missione Aziendale	4
1.1.4 Il presidio di Vimercate	7
1.2 L'U.O.C. Ingegneria Clinica presso l'ASST Vimercate	11
1.2.1 Organizzazione e attività della Struttura	11
1.2.2 L'attività di manutenzione	12
2 La manutenzione dei dispositivi medici	13
2.1 Definizioni	13
2.1.1 Dispositivo medico	13
2.1.2 Manutenzione correttiva e preventiva	13
2.2 Modalità organizzative di manutenzione nelle strutture sanitarie	15
2.3 La manutenzione del parco tecnologico in ASST Vimercate	16
2.3.1 Manutenzioni programmate	18
2.3.2 Manutenzioni correttive	21
2.4 Esempio di manutenzione correttiva: Risonanza magnetica Philips Achieva	24
2.5 Esempio di manutenzione programmata: Ecografo	24
3 Industria 4.0: tra manutenzione predittiva e tecnologie digitali	25
3.1 Industria 4.0 e tecnologie emergenti	25
3.2 Manutenzione predittiva: il contesto normativo	26
3.3 Vantaggi della manutenzione predittiva	27
3.4 Metodi analitici di predizione	30

3.4.1	Classificazione	32
3.4.2	Regressione	35
3.4.3	Detezione di anomalie	39
4	La manutenzione predittiva in Ospedale: l'esperienza di Vimercate	42
4.1	Indagine preliminare	42
4.2	L'analisi dei log macchina	44
4.3	Le lavaendoscopi MEDIVATORS ®ISA ®	46
4.4	Raccolta Dati	50
4.5	Conversione dei file di backup	51
4.6	Modellizzazione	54
4.6.1	Apprendimento ad istanza multipla	55
4.6.2	Text mining: creazione del Corpus	56
4.6.3	Text mining: Document Term Matrix	58
4.6.4	Fase di training e testing	61
4.7	Scelta del modello ottimale	63
5	INSIGHT: IoT al servizio del reparto di endoscopia	66
6	Conclusioni	67
Bibliografia		68

Elenco delle figure

1.1	Rete territoriale ATS Brianza	3
1.2	Organigramma aziendale	5
1.3	Ospedale di Vimercate, ingresso	8
1.4	Ospedale di Vimercate, vista aerea	10
2.1	Ruoli dei diversi attori coinvolti nella manutenzione programmata	19
2.2	Documenti di riferimento relativi alla procedura di manutenzione programmata	19
2.3	Ruoli dei diversi attori coinvolti nella manutenzione correttiva	23
2.4	documenti di riferimento	23
3.1	Numero di prestazioni ambulatoriali in ASST Vimercate per tipologia di esame.	28
3.2	Modello di neurone artificiale.	36
3.3	Topologia di rete neurale feedforward.	38
3.4	Topologia di una LSTM	40
4.1	Lava-Sterilizzatrice MEDIVATORS®ISA®	47
4.2	Report stampato dalla lava-sterilizzatrice MEDIVATORS ®ISA ®	49
4.3	Schema di lavoro	50
4.4	Divisione degli scontrini nelle parti di Intro, Header e Footer	53
4.5	Tabella ottenuta dalla conversione dei file testuali degli scontrini.	54
4.6	Apprendimento ad istanza multipla. La prima borsa è associata ad una etichetta “negativa” in quanto al di fuori dell’intervallo di predizione dei 7 giorni.	56
4.7	Valore dato dalla tf-idf per due termini generici t_1 e t_2 , ipotizzando la stessa frequenza ($f_{d,t_{1,2}} = 3$) all’interno di un generico documento d . t_1 compare in meno documenti del corpus ($n_{t_1} = 4$), di conseguenza il suo peso risulta maggiore.	59
4.8	Frequenza dei 40 termini più comuni nel corpus degli scontrini.	60

Elenco delle tabelle

4.1 Document Term Matrix	61
4.2 Matrice di confusione	63
4.3 Performance dei modelli ottenuti	64

Sommario

Il sommario deve contenere 3 o 4 frasi tratte dall'introduzione di cui la prima inquadra l'area dove si svolge il lavoro (eventualmente la seconda inquadra la sottoarea più specifica del lavoro), la seconda o la terza frase dovrebbe iniziare con le parole "Lo scopo della tesi è..." e infine la terza o quarta frase riassume brevemente l'attività svolta, i risultati ottenuti e deventuali valutazioni di questi.

Abstract

The abstract must contains 3 or 4 sentences from the introduction. The first one should be related to the area of the study, with the second one more, possibly, specific about the same area. The third one should start with the formula: “The goal of this dissertation is...”. Finally, the fourth sentence should be a brief summary of the activity, with the relatives results and possible evaluation of the same.

1 Introduzione e contesto

Questo capitolo introduttivo sarà dedicato all'organizzazione dell'ASST Vimercate, alla sua origine istituzionale e alla descrizione del ruolo e delle attività svolte dalla U.O.C. Ingegneria Clinica, con particolare riferimento al tema della gestione della manutenzione dagli aspetti più convenzionali a quelli più innovativi oggetto del presente elaborato di tesi.

1.1 Organizzazione della ASST Vimercate

Il presente lavoro di tesi è frutto dell'esperienza di tirocinio svolta presso l'ASST Vimercate da Marzo a Ottobre 2019.

L'Azienda Socio Sanitaria Territoriale di Vimercate è parte del Sistema Socio-Sanitario Regionale, nell'ambito del quale esercita le proprie funzioni di prevenzione, diagnosi, cura, riabilitazione e presa in carico, con la finalità primaria di contribuire al miglioramento dello stato di salute della popolazione.

L'attuale assetto organizzativo dell'Azienda deriva dalla riorganizzazione del Sistema Socio-Sanitario Lombardo conseguente all'approvazione della L.R. n. 23 del 11 agosto 2015.

Le ASST si articolano in due settori aziendali rispettivamente definiti rete territoriale e polo ospedaliero.

Il settore aziendale polo ospedaliero si articola in presidi ospedalieri e/o in dipartimenti organizzati in diversi livelli di intensità di cura, tecnologica e organizzativa, in coerenza con la normativa nazionale.

Il settore aziendale rete territoriale eroga prestazioni specialistiche, di prevenzione sanitaria, diagnosi, cura e riabilitazione a media e bassa complessità, le cure intermedie e garantisce le funzioni e le prestazioni medico-legali. Il settore rete territoriale eroga, altresì, le prestazioni distrettuali fornite alle persone, che erano erogate dalle ASL ad eccezione di quelle ora espressamente attribuite alle ATS, tra cui: assistenza

specialistica ambulatoriale, servizi per la prevenzione e la cura delle tossicodipendenze, servizi consultoriali, servizi per disabili, anziani, cronici e fragili, assistenza domiciliare integrata [1].

1.1.1 Polo Ospedaliero

Si riportano le strutture facenti parte del Polo Ospedaliero:

- Ospedale Vittorio Emanuele III a Carate Brianza
- Ospedale C. Borella a Giussano
- Ospedale Trabattoni e Ronzoni a Seregno
- Ospedale Civile a Vimercate
- Ambulatori ospedalieri e territoriali di Arcore, Besana Brianza, Carate Brianza, Giussano, Lentate sul Seveso, Lissone, Macherio, Meda, Seregno, Seveso, Usmate, Vimercate Centri odontostomatologici di Arcore, Besana Brianza, Carate Brianza, Giussano, Lissone, Macherio, Seregno, Vimercate
- Servizi dialisi di Agrate Brianza, Carate Brianza, Lissone, Seregno, Vimercate
- Centri psicosociali/Ambulatori psichiatrici di Carate Brianza, Besana Brianza, Lissone, Seregno, Vimercate Centri psichiatrici diurni/residenziali di Bernareggio, Besana Brianza, Carate Brianza, Meda, Seregno
- Ambulatori territoriali neuropsichiatrie infanzia/adolescenza di Besana Brianza, Giussano, Lentate sul Seveso, Lissone, Seregno, Seveso, Usmate
- Centri neuropsichiatrici infanzia/adolescenza diurni/residenziali di Besana Brianza, Lissone

1.1.2 Rete Territoriale

- Distretto di Carate Brianza
- Distretto di Seregno
- Distretto di Vimercate
- Consultori familiari di Arcore, Carate Brianza, Concorezzo, Giussano, Lentate sul Seveso, Lissone, Meda, Seveso, Seregno, Vimercate

1 INTRODUZIONE E CONTESTO

- Centro riabilitativo diurno di Usmate
- Nuclei operativi alcologia di Seregno, Vimercate
- Servizio dipendenze di Carate Brianza

Nella mappa riportata in figura 1.1 è visibile, in blu, l'intero territorio della ATS Brianza e, in rosso, la rete territoriale della ASST Vimercate. Le puntine indicano i presidi ospedalieri dei rispettivi distretti territoriali.

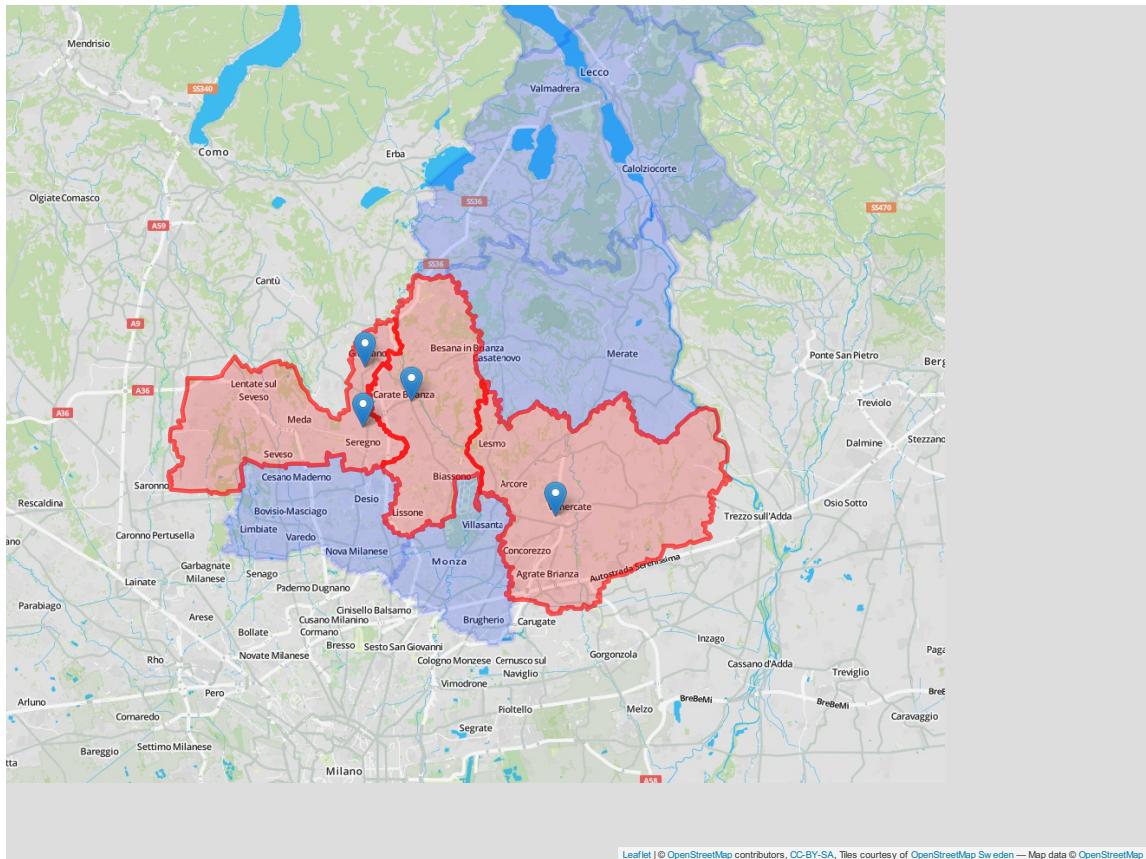


Figura 1.1: Rete territoriale ATS Brianza

1.1.3 La Missione Aziendale

La missione dell’Azienda è quella di contribuire alla promozione, al mantenimento e al recupero delle condizioni di salute della popolazione, erogando prestazioni sia sanitarie sia sociosanitarie, secondo standard di eccellenza, di media e elevata complessità, nel rispetto della dignità della persona, del bisogno di salute, dell’equità dell’accesso all’assistenza, della qualità delle cure e della loro appropriatezza, dell’economicità nell’impiego delle risorse ed in coerenza con i principi e gli obiettivi definiti dal Piano Sanitario Nazionale e dal Piano Sanitario Regionale.

L’Azienda, conformemente alle indicazioni Regionali, ha strutturato un sistema “che si prenda cura” della persona e del suo benessere, collaborando attivamente con tutti gli attori del Sistema Socio-Sanitario, al fine di perseguire l’obiettivo di rafforzare nel cittadino la percezione di un’identità comune, garantendo così a tutti parità di accesso ai servizi.

In figura 1.2 viene riportato l’organigramma aziendale in cui sono evidenziati gli organi di vertice e le strutture (semplici e complesse) cui essi fanno capo.

1 INTRODUZIONE E CONTESTO

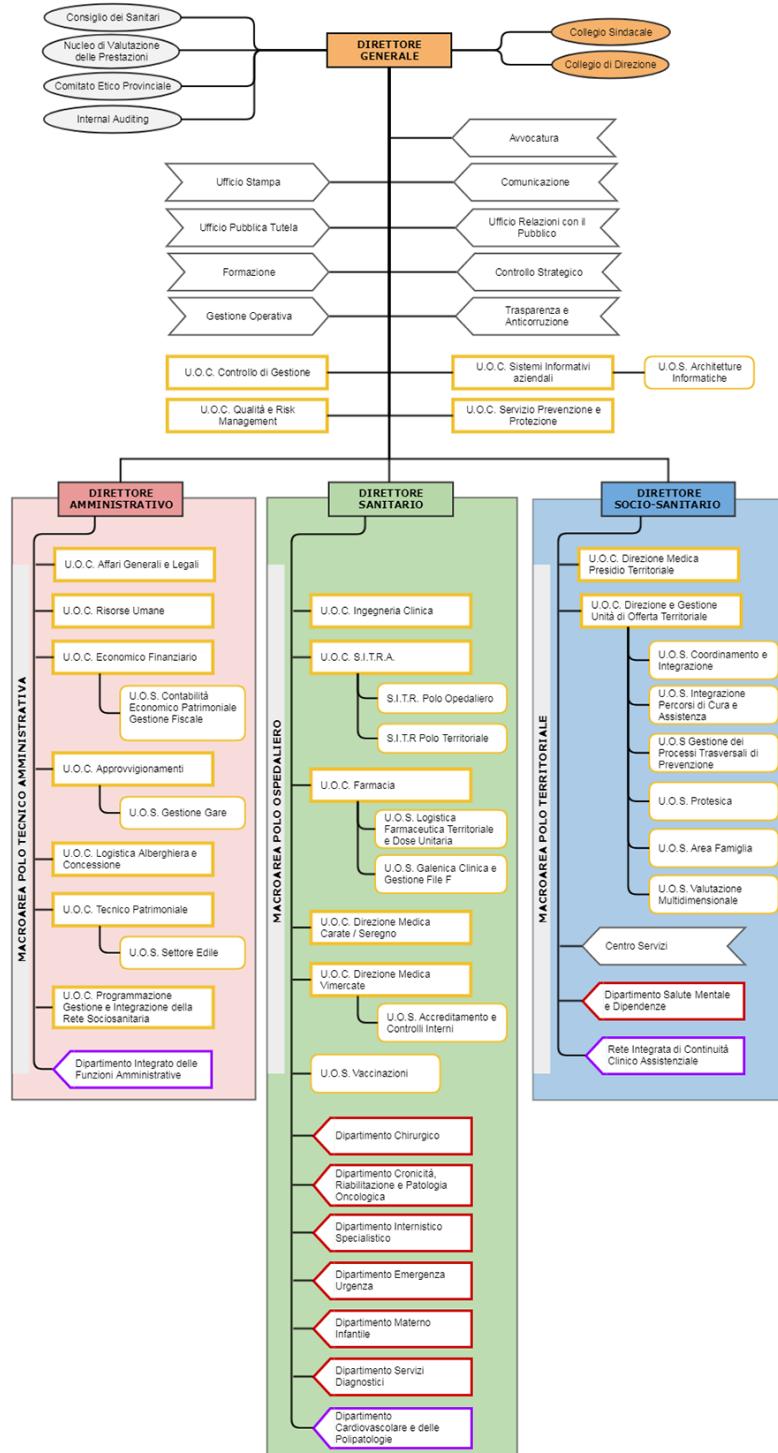


Figura 1.2: Organigramma aziendale

Costituiscono la Direzione Strategica dell’Azienda, unitamente al Direttore Generale, il Direttore Amministrativo, il Direttore Sanitario ed il Direttore Sociosanitario. Essi partecipano con il Direttore Generale alla direzione dell’Azienda, assumono diretta responsabilità delle funzioni attribuite alla loro competenza e concorrono, con proposte e pareri, alla formulazione delle decisioni della Direzione Strategica. Sono, quindi, soggetti attori, unitamente al Direttore Generale, di quella funzione di programmazione, di allocazione delle risorse, di committenza dei servizi e di valutazione dei risultati, propria della Direzione Strategica. Essi esercitano le loro funzioni attraverso il processo di budget, che li vede direttamente coinvolti e responsabili dal momento delle scelte aziendali di programmazione alla negoziazione dei budget con i Direttori dei Dipartimenti ed i Responsabili delle Funzioni di Staff, al monitoraggio e all’eventuale revisione degli obiettivi di budget fissati.

In particolare, le funzioni proprie di ogni Direzione possono essere riassunte come segue:

- **Direzione Generale:** il Direttore Generale rappresenta legalmente l’Azienda ed esercita tutti i poteri di gestione complessiva, coadiuvato dai Direttori Amministrativo, Sanitario e Sociosanitario. Egli ha il compito di combinare i fattori produttivi per il perseguimento degli obiettivi sanitari posti dalla Regione, in condizioni di massima efficienza ed efficacia. Il Direttore Generale è quindi responsabile del governo clinico dell’azienda, con riferimento sia alla qualità, che all’efficacia e all’efficienza tecnica. Il governo clinico si realizza nell’esercizio della responsabilità del miglioramento continuo dei servizi erogati, nella salvaguardia di alti standard assistenziali da parte dei professionisti dell’organizzazione, ed è volto ad assicurare l’adozione di sistemi di garanzia e di qualità, compatibilmente con le risorse disponibili.
- **Direzione Amministrativa:** il Direttore Amministrativo dirige i servizi amministrativi della Azienda Socio Sanitaria Territoriale. In particolare:
 - Svolge attività di direzione e coordinamento, nei confronti dei direttori delle strutture amministrative complesse.

- Svolge attività necessarie per il controllo di gestione dell’azienda ed il controllo di qualità dei servizi come la nomina dei responsabili delle strutture organizzative professionali e funzionali dell’area amministrativa.
 - Persegue il buon andamento e l’imparzialità dell’azione amministrativa, con particolare riferimento agli aspetti giuridico-amministrativi ed economico-finanziari.
- **Direzione Sanitaria:** Il Direttore Sanitario dirige i servizi sanitari ai fini organizzativi ed igienico-sanitari. In particolare:
 - Svolge attività di indirizzo, coordinamento, supporto, verifica nei confronti dei responsabili dei servizi sanitari e promuove l’integrazione dei servizi stessi.
 - Esercita per tutte le strutture ed i Centri di responsabilità sanitari dell’azienda le funzioni di verifica e di controllo dei risultati conseguiti rispetto agli obiettivi assegnati.
 - Coordina le strutture dell’azienda ospedaliera e garantisce l’integrazione tra le stesse, anche attraverso la predisposizione di appositi protocolli.
 - **Direzione Socio-Sanitaria:** il Direttore Sociosanitario dirige il settore aziendale rete territoriale della Azienda Socio Sanitaria Territoriale.

1.1.4 Il presidio di Vimercate

Nella figura 1.3 viene riportata un’immagine dell’ingresso principale del nuovo Ospedale di Vimercate.



Figura 1.3: Ospedale di Vimercate, ingresso

Il Presidio di Vimercate vanta una storia millenaria: nacque, grazie all'azione dei Monaci di Sant'Ambrogio, come ricovero per pellegrini nell'833 con il nome di "L'Hospitale di San Damiano". Fu ribattezzato nel 1447 dapprima con il nome di "Hospitale di San Giuseppe" e quindi "Hospitale Pauperum", infine venne indicato con il nome di "Hospitale del Borgo di Vicomercato", a seguito di un'opera di riorganizzazione indotta dall'allora arcivescovo di Milano, Enrico Rampini. Nel 1778, sotto la spinta del Conte Trottì, sovrintendente dell'ospedale, il nosocomio venne ristrutturato e si costruì un nuovo fabbricato, l'"Opera Pia ed Ospitale dei Poveri e luoghi Pii Uniti di Vimercate" inaugurato nel 1783 e tuttora esistente. Altri interventi di ristrutturazione e di riforma portarono alla costruzione nel 1926 dell'Ospedale di Circolo, nel 1937 del Padiglione Sanitario, nel 1954 del Dispensario, nel 1958 del Padiglione Maternità, e nel 1961 del Poliambulatorio Inam. Nel 1973 venne inaugurato il primo monoblocco chirurgico: Vimercate diventò il presidio ospedaliero della Ussl 60, e nel 1995 della Ussl 28.

1 INTRODUZIONE E CONTESTO

Dal gennaio 1998 Vimercate è l’Ospedale capofila dell’Azienda Ospedaliera “Ospedale Civile di Vimercate”, ridenominata nell’anno 2009 “Azienda Ospedaliera di Desio e Vimercate”.

Importante è sottolineare che dal 30 ottobre 2010 l’Ospedale di Vimercate ha iniziato una nuova vita: l’attività si è infatti trasferita nella nuova sede di via Santi Cosma e Damiano, nel Nuovo Complesso Ospedaliero voluto, insieme a Como, Legnano, Milano Niguarda e Bergamo, dalla Giunta Regionale lombarda.

Il nuovo Ospedale, vero fiore all’occhiello dell’Azienda, progettato dall’arch. Mario Botta e costruito in soli tre anni, si sviluppa su tre piani fuori terra e due interrati, per una superficie totale di 113.000 mq (nella figura 1.4 si riporta una ricostruzione di una visualizzazione aerea tratta dai piani di costruzione del nuovo Ospedale [2]).

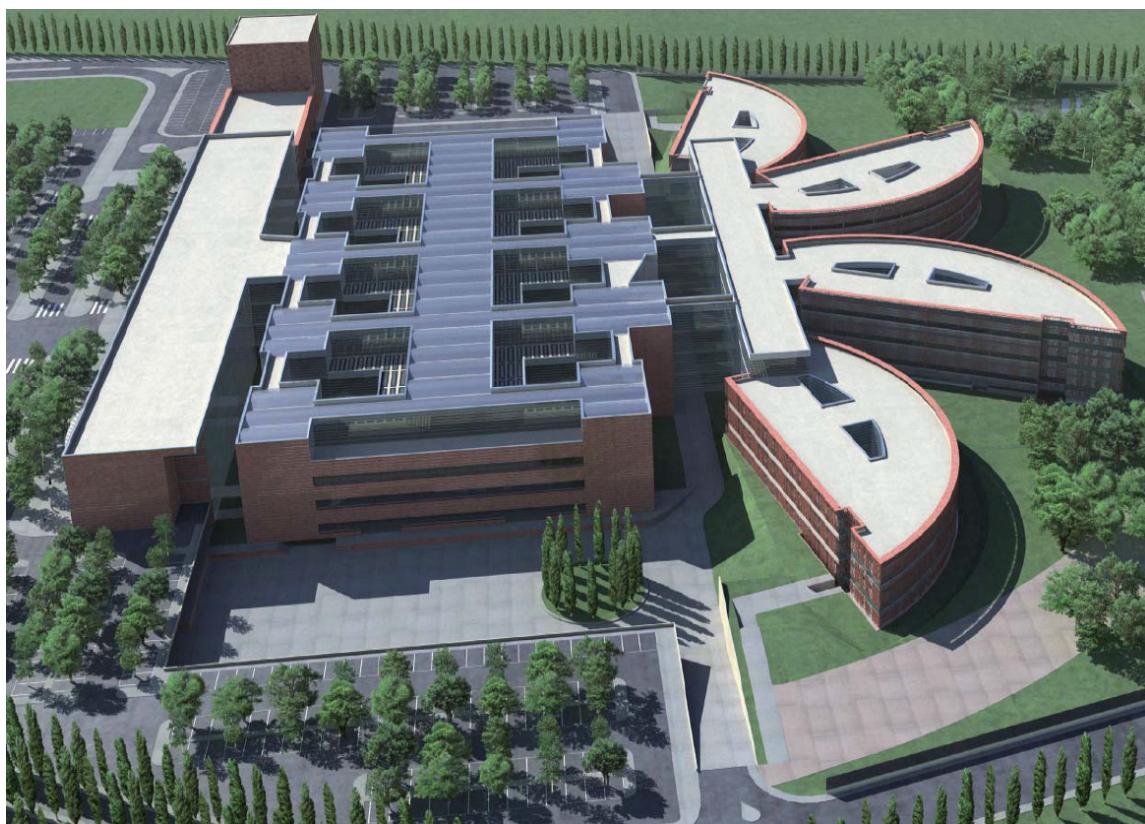


Figura 1.4: Ospedale di Vimercate, vista aerea

L'architettura, la tecnologia e l'organizzazione sono tutte finalizzate a porre al centro dei percorsi logistici e diagnostico-terapeutici il paziente, con le sue patologie e il suo vissuto familiare e sociale; anche per questo motivo l'assistenza è organizzata secondo il modello per intensità di cura.

1.2 L’U.O.C. Ingegneria Clinica presso l’ASST Vimercate

1.2.1 Organizzazione e attività della Struttura

L’Unità Operativa Complessa (U.O.C.) Ingegneria Clinica, in staff alla Direzione Sanitaria Aziendale come da organigramma aziendale in figura 1.2, partecipa alla cura della salute garantendo un uso sicuro, appropriato, efficiente ed economico delle apparecchiature elettromedicali in dotazione all’Azienda. L’U.O.C. Ingegneria Clinica si articola territorialmente con una sede centrale a Vimercate e un ufficio a Carate, è diretta da un Ingegnere Biomedico e si avvale delle seguenti risorse umane:

- 3 collaboratori tecnici
- 2 assistenti tecnici
- 1 collaboratore amministrativo

La Struttura si occupa dell’aspetto gestionale e strategico delle tecnologie biomediche, svolgendo un ruolo di supporto alla Direzione Strategica attraverso:

- La programmazione e la valutazione degli acquisti
- La manutenzione delle apparecchiature elettromedicali in dotazione all’Azienda
- Valutazioni di Hospital Based Health Technology Assessment finalizzate all’introduzione di nuove tecnologie ed ottimizzazione nell’utilizzo delle risorse tecnologiche a disposizione.

La Struttura svolge la propria attività coordinandosi con altre strutture tecniche, amministrative, con il personale sanitario dirigente e del comparto per il perseguimento di obiettivi comuni e la realizzazione di attività e progetti di interesse aziendale. Le principali attività svolte all’interno dell’U.O.C. Ingegneria Clinica comprendono:

- **Gestione strategica degli investimenti:** effettuata mediante azioni di riconoscimento delle esigenze (in supporto alla Direzione Strategica), attività di programmazione delle acquisizioni di apparecchiature elettromedicali attraverso strategie di Health Technology Assessment, attività di supporto alla U.O.C. Approvvigionamenti per la predisposizione di capitolati di fornitura tramite elab-

borazione di fabbisogni e dettagli tecnici necessari e infine attività di valutazione tecnica delle offerte pervenute e partecipazione alle commissioni giudicatrici.

- **Gestione tecnica e operativa:** include la gestione e la supervisione delle attività facenti parte del “ciclo di vita” delle tecnologie biomedicali (accettazione e collaudo, manutenzione correttiva, preventiva, straordinaria, verifiche elettriche, controlli qualità, gestione database tecnico delle apparecchiature, training degli operatori, dismissione) tramite il supporto di un servizio di manutenzione esternalizzata e contratti con i produttori.
- **Gestione dei contratti:** le specifiche competenze sono relative all’esecuzione dei contratti di manutenzione, acquisizione in noleggio o proprietà di apparecchiature elettromedicali.

1.2.2 L’attività di manutenzione

La gestione delle attività di manutenzione, come visto nel precedente paragrafo, rientra in una delle attività core del Servizio di Ingegneria Clinica (SIC), ovvero la gestione tecnica e operativa. In particolare, l’Azienda ha adottato la scelta di esternalizzare il servizio di manutenzione delle apparecchiature elettromedicali. A seguito di procedura di gara per l’affidamento del servizio, dal 1 dicembre 2015 il servizio di manutenzione esternalizzata viene erogato dalla società Tecnologie Sanitarie S.p.A. (comunemente e brevemente chiamata “Global Service”) che si occupa appunto delle attività di manutenzione (correttiva, preventiva, verifiche elettriche, controlli qualità) secondo quanto stabilito dal Capitolato di gara di appalto e nel contratto di affidamento.

2 La manutenzione dei dispositivi medici

Nella prima parte di questo capitolo verranno introdotte le definizioni di “dispositivo medico”, “manutenzione correttiva” e “manutenzione preventiva” secondo le norme attualmente in vigore. Successivamente verranno illustrate le principali strategie di manutenzione negli ospedali, con riferimento alla realtà ospedaliera di Vimercate. Verranno quindi presentati gli attori coinvolti nella gestione della manutenzione e i loro ruoli. Infine verranno riportati alcuni esempi di manutenzione (sia correttiva che preventiva) a cui si è assistito durante il periodo di tirocinio presso l’ospedale.

2.1 Definizioni

2.1.1 Dispositivo medico

Un dispositivo medico, secondo quanto stabilito dal decreto legislativo 24/02/1997 nr. 46, con la quale viene recepita, in Italia, la direttiva europea 93/42/EEC (MDD, medical device directive), viene definito come:

“qualsiasi strumento, apparecchio, impianto, sostanza o altro prodotto, utilizzato da solo o in combinazione, compreso il software informatico impiegato per il corretto funzionamento, e destinato dal fabbricante ad essere impiegato nell'uomo a scopo di diagnosi, prevenzione, controllo, terapia o attenuazione di una malattia; di diagnosi, controllo, terapia, attenuazione o compensazione di una ferita o di un handicap; di studio, sostituzione o modifica dell'anatomia o di un processo fisiologico; di intervento sul concepimento, il quale prodotto non eserciti l'azione principale, nel o sul corpo umano, cui è destinato, con mezzi farmacologici o immunologici né mediante processo metabolico ma la cui funzione possa essere coadiuvata da tali mezzi.[3]”

2.1.2 Manutenzione correttiva e preventiva

La manutenzione correttiva viene definita come :

“manutenzione eseguita a seguito della rilevazione di una avaria e volta a riportare un’entità nello stato in cui essa possa eseguire una funzione richiesta.”(CEI 62-122).

Questa strategia consiste quindi nel mantenere la macchina in esercizio fintanto che il manifestarsi o il progredire di una anomalia costringe il gestore a fermare la macchina. La strategia, presenta sia aspetti positivi che negativi. Il fattore positivo è un costo di manutenzione e di fermo macchina pressoché nullo fin tanto che la macchina funziona come dovrebbe. I fattori negativi possono invece essere sintetizzati in:

- elevata perdita di ricavi dovuti al fermo macchina per guasto
- imprevedibilità dell’intervento e quindi delle eventuali operazioni di deviazione del flusso produttivo in corso
- generalmente elevati costi di riparazione. Un guasto che si protrae per molto tempo ad un componente può danneggiare altri dispositivi della macchina e può renderla inutilizzabile definitivamente.

La manutenzione preventiva (generica) viene definita come:

“manutenzione eseguita ad intervalli regolari predeterminati o in accordo a criteri prescritti e volta a ridurre la probabilità di guasto o la degradazione del funzionamento di un dispositivo medico.”(CEI 62-122 – UNI 9910).

La programmabilità dell’intervento consente una maggiore organizzazione dei lavori di manutenzione e garantisce la possibilità di gestire i fermo macchina nella maniera più conveniente. Vengono inoltre date le definizioni (CEI 62-122) di:

- **Manutenzione preventiva di primo livello:** *“controlli preventivi che dovrebbero essere effettuati dall’operatore consistenti nell’ispezione a vista e, se applicabili in semplici prove utilizzando i dispositivi medici. Queste prove/controlli devono essere descritti nel manuale d’uso.”*
- **Manutenzione preventiva di secondo livello:** *“manutenzione preventiva effettuata da personale qualificato, consistente nell’ispezione visiva, nelle misure dei parametri importanti ai fini della sicurezza, nell’accertamento che la man-*

tenzione di primo livello sia stata correttamente eseguita, nonché nell'esecuzione dei programmi di manutenzione prescritti dal costruttore.”

2.2 Modalità organizzative di manutenzione nelle strutture sanitarie

La gestione e l'attuazione della manutenzione delle varie apparecchiature elettromedicali presenti in una struttura sanitaria, sono attività core del Servizio di Ingegneria Clinica, che possono essere riassunte con il termine “Techonology Management”. Esso ha come obiettivo la soddisfazione di un bisogno primario comune a tutte le strutture sanitarie, ovvero l'ottimale gestione e l'utilizzo sicuro e appropriato di un parco macchine che, sviluppandosi di pari passo con lo sviluppo tecnologico, diventa sempre più diversificato e complesso.

In questo scenario di continua innovazione e crescente complessità tecnologica dei macchinari, si capisce la trasformazione che ha interessato negli ultimi anni l'attività manutentiva. Essa, concepita negli anni passati esclusivamente in funzione del ripristino di apparecchiature mediche non funzionanti, sta progressivamente mutando verso una vera e propria funzione manageriale volta, in ultima analisi, al miglioramento della qualità dell'assistenza al paziente del cui iter diagnostico-terapeutico l'elemento tecnologico gioca un ruolo fondamentale. Questo miglioramento di qualità è ottenuto grazie alla risoluzione degli obiettivi cardine dell'attività manutentiva, volti alla riduzione dei rischi connessi all'uso dei dispositivi medici, a diminuire i tempi di inutilizzo, a prevenire i guasti, a garantire la qualità delle prestazioni erogate, ad ottimizzare quindi, la durata fisiologica del prodotto.

Nelle diverse strutture sanitarie in cui è stato organizzato un Servizio di Ingegneria Clinica (SIC), è possibile individuare tre diversi modelli di organizzativi “ideali”:

- **SIC interno:** istituito internamente alla struttura sanitaria e con personale dipendente dalla struttura stessa. La tipologia di un Servizio di Ingegneria Clinica totalmente interno all'Azienda Sanitaria prevede che tutte le funzioni proprie dell'ingegneria clinica siano svolte con personale interno dipendente. La gestione della tecnologia secondo quest'ottica si basa sul presupposto che il personale sia

altamente qualificato, che conosca dettagliatamente la situazione dell'ente, che abbia solidi legami con l'organizzazione sanitaria, tecnica ed amministrativa.

- **SIC esterno:** istituito mediante appalto di servizi ad una società di servizi esterna che operi in modo coordinato e continuativo con l'ente. Questa strategia consiste nell'affidamento del servizio globale di manutenzione ad un'azienda esterna o ad un consorzio di aziende, attraverso il quale la struttura sanitaria delega completamente la manutenzione (preventiva, correttiva, prove di sicurezza, controlli di qualità e tutte le attività tipiche di un ufficio tecnico) ad un'azienda esterna ad un prezzo complessivo stabilito solitamente attraverso il metodo della gara.
- **SIC misto:** dove solo alcune attività sono affidate all'interno mentre altre sono affidate a una società di servizi esterna che si integra nella struttura sanitaria. Questa tipologia rappresenta una modalità intermedia di organizzazione che prevede l'affidamento ad una società di servizi di solo una parte delle funzioni di ingegneria clinica (nella maggior parte dei casi l'effettuazione delle verifiche periodiche di sicurezza elettrica e/o la manutenzione), mentre le restanti sono svolte da personale interno.

2.3 La manutenzione del parco tecnologico in ASST Vimercate

L'ultima modalità organizzativa citata, come preannunciato nel capitolo precedente (1.2.2), è quella adottata dalla ASST Vimercate. Il Servizio di gestione delle apparecchiature mediche prevede in particolare:

- l'effettuazione delle manutenzioni programmate e delle manutenzioni correttive
- le verifiche di accettazione e collaudo
- la predisposizione e l'aggiornamento del piano di manutenzione

Prima di descrivere nel dettaglio le attività di manutenzione (correttiva e programmata) sopracitate, è necessario specificare le forme di presenza con le quali una tecnologia biomedica può essere presente in una struttura sanitaria, dalle quali dipende l'or-

ganizzazione della manutenzione per la tecnologia. Le forme di presenza sono così definite:

DA RIVEDERE

- Acquisto: l'apparecchiatura acquistata è di proprietà dell'azienda ospedaliera a seguito del pagamento corrisposto al fornitore.
- Noleggio: l'apparecchiatura in noleggio non è di proprietà dell'azienda ospedaliera. L'azienda versa un canone di noleggio all'azienda fornitrice per il periodo di tempo in cui il dispositivo sarà effettivamente utilizzato.
- Service: anche in questo caso l'apparecchiatura non è di proprietà dell'azienda ospedaliera. L'azienda fornitrice si impegna a mettere a disposizione all'azienda ospedaliera il dispositivo in questione a titolo gratuito per l'intero periodo di validità del contratto stipulato.
- Comodato d'uso

DA RIVEDERE

Tutte queste forme di presenza sono caratterizzate da una gestione della manutenzione simile tra loro, direttamente collegata al periodo di garanzia della macchina. Fino allo scadere della garanzia, infatti, l'azienda fornitrice del macchinario è tenuta a comunicare al servizio di ingegneria clinica un calendario (il piano di manutenzione) in cui vengono specificate le date di intervento che seguiranno il collaudo del macchinario stesso. Nello specifico, in questo calendario sono presenti:

- gli interventi di manutenzione ordinaria che l'azienda fornitrice svolgerà (di numero e frequenza variabile a seconda della tecnologia),
- eventuali interventi di manutenzione correttiva,
- verifiche elettriche
- controlli periodici di qualità (calibrazioni ecc.)

Una volta terminato il periodo di garanzia (ovviamente inferiore rispetto alla vita utile effettiva del macchinario), sarà compito del Global Service effettuare tutte le attività di manutenzione espresse precedentemente, tramite le competenze del personale tecnico specializzato (ingegneri biomedici e tecnici informatici ed elettronici).

2.3.1 Manutenzioni programmate

Per quanto riguarda le manutenzioni programmate, il Global si avvale di schede tecniche, specifiche per ogni classe di dispositivo, che contengono procedure guidate sulle operazioni da seguire in sede di manutenzione programmata. Queste schede tecniche sono redatte unendo sia le indicazioni del fabbricante del dispositivo, sia l'esperienza tecnica del Global Service, ottenendo quindi una procedura mista. Un esempio di procedura utilizzata dal Global per condurre una manutenzione programmata viene riportata nel capitolo 2.5

Di seguito si riporta quanto presente nella Procedura Operativa Aziendale dal titolo “Manutenzione programmata” che ha lo scopo di descrivere le modalità con cui l’UOC Ingegneria Clinica gestisce le manutenzioni programmate di apparecchiature elettromedicali (indicate d’ora in avanti con la lettera “a”) nell’Azienda Ospedaliera al fine di garantirne l’integrità fisico-funzionale e l’affidabilità.

I ruoli delle figure coinvolte nella procedura di manutenzione programmata vengono riportati in figura 2.1. La relativa documentazione di riferimento è riportata in figura 2.2.

2 LA MANUTENZIONE DEI DISPOSITIVI MEDICI

SIC	Approva i piani DOC_CPM_GS di manutenzione stesi da GS Supervisiona l'attività manutentiva di GS, avviando le pratiche di penali qualora i tempi e le modalità di intervento non siano rispettati secondo il CSA
EST	Esegue le attività manutentive sulle (a) su richiesta del GS, e su tutte le (a) in contratto, comodato, noleggio, leasing, service Emette DOC_PM_EST a conclusione dell'intervento manutentivo
SS/SC	Firma il DOC_PM_EST e il DOC_PM_RIASS
GS	Inserisce i piani di manutenzione DOC_CPM_GS, approvati da SIC, nel software di gestione delle (a)
	Esegue le manutenzioni programmate sulle (a) servendosi dei DOC_PM_GS, avvisando in anticipo le SS/SC, prima dell'intervento Applica le ETICH_PM_GS sulle (a) su cui vengono eseguite le PM Genera un DOC_OdL_GS in corrispondenza del DOC_PM_EST o del DOC_PM_GS Monitora le manutenzioni programmate sulle (a) in garanzia, comodato, noleggio, service (affidate ad EST) sollecitandole se necessario Aggiorna sul software di gestione delle (a) lo stato delle manutenzioni ed archivia i DOC_OdL_GS Predisponde e firma DOC_PM_RIASS

Figura 2.1: Ruoli dei diversi attori coinvolti nella manutenzione programmata

<u>Documenti di Riferimento</u>	
DOC_CPM_GS	Calendario periodico delle manutenzioni programmata per le (a)
DOC_PM_GS	Schede/Protocolli di manutenzione programmata
DOC_PM_EST	Verbale di manutenzione programmata di (a) effettuata da EST
DOC_PM_RIASS	Elenco cumulativo delle (a) oggetto di manutenzione programmata in un periodo
DOC_OdL_GS	Scheda di intervento a cui viene associata una manutenzione programmata di una (a) da parte di GS o di EST
ETICH_PM_GS	Etichetta apposta da GS su (a) a giustificazione manutenzione programmata effettuata

Figura 2.2: Documenti di riferimento relativi alla procedura di manutenzione programmata

Sono responsabili della procedura operativa di manutenzione programmata di una (a) le seguenti figure:

- Gli utilizzatori delle varie SS/SC che opportunamente informati dal fornitore della (a) al momento della consegna e del collaudo e, supportati poi dai tecnici del Global Service (GS) durante l'utilizzo, eseguono ordinariamente tutti gli interventi previsti di manutenzione/pulizia/controllo giornaliero per il corretto mantenimento della stessa;
- Il GS che secondo i piani ed i calendari stabiliti dal Capitolato Speciale d'Appalto (CSA), esegue periodicamente e nel rispetto della normativa, le verifiche di sicurezza generale, le manutenzioni preventive, i controlli prestazionali/manutenzioni particolari, i controlli di qualità e le tarature su tutte le (a) in uso, supportando le ditte esterne (EST) per le apparecchiature non proprietarie, in garanzia e sotto contratto a cui è affidata nello specifico l'attività manutentiva;
- Il SIC che, una volta approvati i piani di manutenzione delle (a) stesi dal GS in virtù del CSA, controlla il corretto svolgimento delle attività programmate

Su specifiche classi di (a) la manutenzione programmata viene completata con l'attività di taratura. Per taratura si intende “ la verifica ed il controllo delle prestazioni metrologiche di una (a)” ; scopo della taratura è accertarsi che il livello di sicurezza e delle prestazioni erogate da una (a) siano all'interno dei limiti definiti dalle norme e dalle tolleranze definite dal fabbricante. In caso di mancato rispetto, la taratura consente di allineare i valori misurati con quelli di riferimento. La taratura, per le (a) proprietarie dell'A.O., è affidata al GS che conserva la strumentazione di misura necessaria per lo svolgimento dell'attività presso il laboratorio tecnico. A taratura avvenuta il tecnico esecutore deve riportare sul DOC_ PM_ GS tutte le specifiche degli strumenti di misura utilizzati (tipologia, marca, modello, s/n ed identificativo). Su ciascuno strumento è applicata un'etichetta che riporta l'ultima taratura avvenuta e la data di scadenza. E' responsabilità di GS inviare presso i centri di calibrazione autorizzati le strumentazioni di misura oggetto di taratura. La taratura sulle (a) proprietarie in A.O. viene eseguita secondo la periodicità stabilita dal CSA [4].

2.3.2 Manutenzioni correttive

Le manutenzioni correttive avvengono in seguito all'insorgenza di un guasto ad un macchinario. Il reparto interessato dal problema compila una richiesta di intervento tramite il software del Global Service, "Coswin8i", descrivendo la natura del guasto e l'urgenza di un intervento correttivo. La "chiamata" viene ricevuta dai tecnici del Global che, se l'urgenza lo permette, pianificano una data in cui eseguire l'azione correttiva, in accordo anche con il reparto. Se invece il guasto richiede un intervento tempestivo, una squadra (solitamente composta da 2 o 3 tecnici) si reca nel reparto di ubicazione del macchinario di interesse e procede, dove possibile, ad effettuare una prima riparazione. Se la problematicità lo richiede, il macchinario viene portato nel laboratorio del Global Service, in cui è possibile effettuare riparazioni sia di tipo elettrico/elettronico, sia di tipo informatico e meccanico.

Nel caso in cui il Global Service non riesca, con mezzi propri, a riparare un eventuale guasto, esso si assume la responsabilità (finanziaria) di contattare il fornitore dell'apparecchio e procede in una delle seguenti modalità:

1. Il Global Service invia l'apparecchio in assistenza al fornitore e comunica al reparto un prospetto delle tempistiche necessarie per la risoluzione del guasto.
2. Il Global Service richiede alla ditta produttrice del dispositivo l'intervento di un tecnico specializzato e stabilisce con essa una data di intervento utile a minimizzare i disagi che il fermo macchina può generare al reparto di interesse.

Per completezza, viene riportato quanto scritto nella Procedura Operativa Aziendale dal titolo "Manutenzione correttiva" che ha lo scopo di descrivere le modalità con cui l'UOC Ingegneria Clinica gestisce le manutenzioni correttive di (a) nell'Azienda Ospedaliera al fine di garantirne l'integrità fisico-funzionale e l'affidabilità.

Sono responsabili della procedura di manutenzione correttiva per una (a) le seguenti figure:

- Le Unità Operative (Coordinatrice e suo delegato) che in caso di mal funzionamento di una (a) deve attivare il GS, tramite richiesta via informatica o

- telefonica, per l'assistenza tecnica del caso;
- Il GS che durante la riparazione di una (a) osserva nel pieno rispetto le prescrizioni della normative vigenti in materia, anche in ordine agli obblighi verso il fabbricante per consentire, a quest'ultimo, la sorveglianza post-vendita e la rintracciabilità delle (a) in ogni momento;
 - Il costruttore che rimane a capo della piena responsabilità per eventuali vizi o difetti originati dalla (a) qualora gli utilizzatori dimostrano di avere assunto tutti i comportamenti atti a preservare lo stato, l'efficienza e il funzionamento della (a) senza alternarne la destinazione d'uso;
 - Il SIC che monitora l'avanzamento di un' attività di riparazione sulla (a) attraverso l'utilizzo del software di gestione Coswin8i, verificando che la risoluzione in un guasto avvenga nel rispetto di quanto descritto nel CSA.

La manutenzione correttiva viene eseguita su tutte le (a) dal GS o dalla EST, o sua emanatrice, su richiesta dell'Unità Operativa utilizzatrice che richiede l'assistenza, oppure ogniqualvolta che durante l'attività di manutenzione preventiva, il GS ravvisi la necessità di una riparazione [5].

Il coinvolgimento delle diverse strutture operative nella manutenzione correttiva di una (a) è descritto nella figura 2.3. La relativa documentazione di riferimento è riportata in figura 2.4.

2 LA MANUTENZIONE DEI DISPOSITIVI MEDICI

SIC	Supervisiona l'attività correttiva di GS, avviando i solleciti e le pratiche di penali qualora i tempi e le modalità di intervento non vengano rispettati in virtù del CSA
	Emette DOC_INT_SIC qualora intende chiedere assistenza alla EST
	Valuta le comunicazioni di irreparabilità da parte della EST per le apparecchiature proprietarie, quindi la dismissione
EST	Esegue le attività manutentive di riparazione sulle (a), in loco o in sede, su richiesta del SIC, su richiesta del GS, e su tutte le (a) non proprietarie, (in contratto, comodato, noleggio, leasing, service..)
	Emette un DOC_CORR_EST
UO	Contatta GS telefonicamente o tramite software di gestione delle (a) per richiedere assistenza su (a)
	Firma DOC_CORR_GS e DOC_CORR_EST
GS	Riceve una richiesta di intervento per una (a) guasta dall'Unità Operativa, cui viene generato, tramite software di gestione delle (a), un OdL identificabile da un n. progressivo
	Svolge l'intervento di manutenzione correttiva su (a), altrimenti, richiede Assistenza alla EST, emettendo DOC_INT_GS
	Chiude l'intervento di manutenzione correttiva emettendo, a completamento DOC_CORR_GS; aggiorna lo stato dell'OdL nel software di gestione delle (a) e archivia la pratica.

Figura 2.3: Ruoli dei diversi attori coinvolti nella manutenzione correttiva

<u>Documenti di Riferimento</u>	
OdL	Ordine/Buono di Lavoro
DOC_CORR_GS	Verbale di manutenzione correttiva per (a) effettuata da Global Service
DOC_INT_SIC	Richiesta Assistenza Tecnica/Preventivo/Consuntivo spese manutenzione correttiva di (a) da SIC alla Ditta Esterna
DOC_INT_GS	Richiesta Assistenza Tecnica/Preventivo/Consuntivo spesa manutenzione correttiva di (a) da GS alla Ditta Esterna
DOC_CORR_EST	Verbale di manutenzione correttiva per (a) effettuata da Ditta Esterna
CSA	Capitolato Speciale d'Appalto Apparecchiature Elettromedicali

Figura 2.4: documenti di riferimento

2.4 Esempio di manutenzione correttiva: Risonanza magnetica Philips Achieva

2.5 Esempio di manutenzione programmata: Ecografo

3 Industria 4.0: tra manutenzione predittiva e tecnologie digitali

In questo capitolo verrà data, dopo una breve introduzione storica, la definizione di “industria 4.0” e di “manutenzione predittiva”. Verranno inoltre riportati i vantaggi nell’utilizzo delle tecnologie digitali a scopo predittivo nell’ottica della manutenzione e verranno descritti i principi teorici dell’analisi predittiva, specificando le principali tecniche di analisi e modelli di machine learning utilizzati attualmente.

3.1 Industria 4.0 e tecnologie emergenti

La storia della manutenzione predittiva è intrinsecamente legata a quella della quarta rivoluzione industriale la quale, a sua volta, si configura come uno sviluppo della terza rivoluzione industriale, definita come rivoluzione digitale. Quest’ultima, iniziata negli anni 80 del secolo scorso, è caratterizzata dalle innovazioni tecnologiche che hanno permesso il “salto tecnologico” dalle tecnologie analogiche e dei dispositivi meccanici alle attuali tecnologie digitali, come ad esempio il pc (personal computer), internet e in generale la branca degli argomenti di interesse della ICT (information and communications technology). In questo contesto, la quarta rivoluzione industriale nasce proprio da queste innovazioni tecnologiche ed è caratterizzata dall’evoluzione di tecnologie emergenti il cui impatto sulla società e sulla qualità della vita non ha eguali in tutta la storia umana. Tra le tecnologie emergenti rientrano la robotica, la nanotecnologia, i computer quantistici, la medicina rigenerativa, l’Industrial Internet of Things, la domotica e l’intelligenza artificiale applicata in svariati campi (per esempio automazione industriale, diagnostica per immagini, business intelligence e analisi di big data).

Con questa premessa, si intuisce come la manutenzione predittiva sia diretta conseguenza di un’ applicazione sinergica delle sopracitate tecnologie. Essa si appropria infatti di metodiche caratteristiche di diversi campi al fine di determinare lo stato di

salute di una tecnologia per prevedere l'istante temporale ottimale in cui condurre le operazioni di manutenzione e quindi il tempo residuo prima di un guasto.

Tra queste metodiche rientra ad esempio l'utilizzo di tecnologie IoT: infatti la valutazione dello stato di salute di un ipotetico parco macchine viene effettuata tramite l'utilizzo di una rete di sensori in grado di comunicare l'andamento nel tempo di alcune variabili di interesse (monitoraggio online). Oppure ancora, come il nome stesso suggerisce, la componente “predittiva” è affidata a più o meno sofisticati, a seconda del contesto, algoritmi di artificial intelligence basati a loro volta sull'applicazione di tecniche di machine learning in grado, in questo caso, di analizzare e predire l'evoluzione di serie temporali sia in modo semi-automatico (apprendimento supervisionato) sia in modo totalmente automatico (apprendimento non supervisionato).

3.2 Manutenzione predittiva: il contesto normativo

Dal punto di vista normativo, la definizione di manutenzione predittiva viene delineata, a livello europeo, nella EN 13306 dove, nella versione attualmente in vigore (EN 13306:2017) essa viene definita come:

“Condition based maintenance carried out following a forecast derived from repeated analysis or known characteristics and evaluation of the significant parameters of the degradation of the item”[6].

La stessa viene recepita in Italia con la UNI EN 13306:2018, secondo la quale per “manutenzione predittiva” si intende:

“Manutenzione su condizione eseguita in seguito a una previsione derivata dall’analisi ripetuta o da caratteristiche note e dalla valutazione dei parametri significativi afferenti il degrado dell’entità”[7].

Dove, sempre secondo la stessa norma, la manutenzione su condizione è definita come:

“Manutenzione preventiva che comprende la valutazione delle condizioni fisiche, l’analisi e le possibili azioni di manutenzione conseguenti”.

La “valutazione”, sempre secondo la sopracitata norma, può avvenire mediante diverse modalità tra le quali:

- Osservazione dell’operatore
- Ispezione
- Collaudo
- Monitoraggio delle condizioni dei parametri del sistema

Tutte queste modalità vengono intese come “svolte secondo un programma, su richiesta o in continuo”.

Riassumendo, quindi, la manutenzione predittiva si configura come un caso “avanzato” di manutenzione preventiva, che mira alla minimizzazione dei tempi di fermo macchina grazie all’applicazione di analisi predittive, con lo scopo di predire, con una certa accuratezza, il tempo rimanente prima di un successivo “guasto” della macchina in esame.

3.3 Vantaggi della manutenzione predittiva

I vantaggi di questa strategia di manutenzione possono essere sintetizzati in quattro punti fondamentali:

1. **Riduzione del tempo di fermo macchina:** Sicuramente il più concreto dei vantaggi della manutenzione predittiva. Abbattendo le probabilità di guasto di un macchinario, si riducono ovviamente i tempi di fermo macchina. I vantaggi di questo risultato sono immediatamente chiari: da una parte si risparmia in tempo e denaro, dall’altra si ha la sicurezza di erogare un servizio in modo continuativo, riducendo quindi la *customer dissatisfaction*, che in sanità è legata alla qualità della cura clinica ricevuta, disponibilità dei servizi e tempi di attesa. La gravità di un fermo macchina è ben intuibile se si considera la quantità di esami che vengono effettuati per mezzo di quella famiglia di macchinari “ad alta incidenza” come TAC, risonanze ed in generale tutti i dispositivi di diagnostica per immagini. In figura 3.1 sono rappresentati le 15 prestazioni ambulatoriali

con il maggior numero di erogazioni in un anno nella ASST di Vimercate in riferimento alle tecnologie sopracitate.

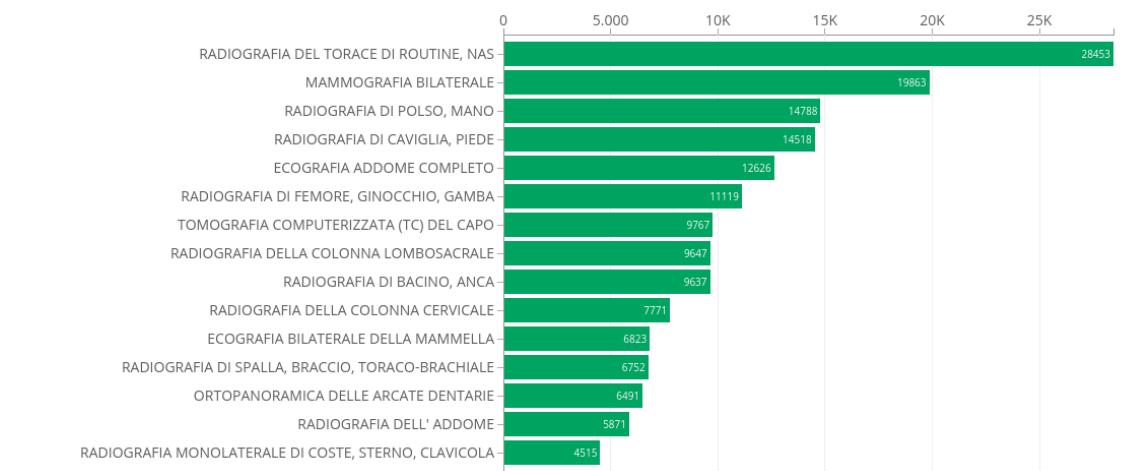


Figura 3.1: Numero di prestazioni ambulatoriali in ASST Vimercate per tipologia di esame.

I dati utilizzati per questa rappresentazione sono stati estratti dal portale di Regione Lombardia adibito alla pubblicazione e consultazione di open data [8]. Appare ben chiaro che i fermo macchina relativi ad apparecchiature di questo tipo, che svolgono mediamente più di dieci mila esami all'anno (si parla quindi di circa trenta esami al giorno), siano delle criticità notevoli ed è facile immaginare la difficoltà con cui un reparto debba provvedere alla riorganizzazione dei crono-programmi in seguito ad un fermo macchina. Obiettivo della manutenzione predittiva è proprio quello di minimizzare queste criticità e, conseguentemente, di ottimizzare la gestione delle tecnologie.

2. **Ottimizzazione delle risorse:** In termini di componenti di un macchinario, tramite applicazioni di manutenzione predittiva è possibile eseguire delle operazioni di sostituzione di componenti prima che questi si degradino oltre un certo stato, andando quindi ad allungare la vita utile dell'apparecchiatura di cui essi fanno parte.

3. Controllo: La raccolta di dati utili all’implementazione di un programma di manutenzione predittiva è un passaggio cruciale per l’ottenimento di un completo controllo del processo che si vuole monitorare. L’obiettivo è quindi quello di conoscere, in ogni momento, lo stato di funzionamento di un macchinario al fine di pianificare e gestire in anticipo situazioni critiche di infungibilità. In particolare, si fa riferimento a tre tipologie di dati:

- Dati ambientali come temperatura, umidità, frequenza delle vibrazioni, ecc.
- Dati storici come informazioni sui guasti passati e operazioni di manutenzione svolte.
- Dati operativi come, ad esempio, informazioni circa l’effettivo utilizzo della macchina

4. Sviluppo continuo e cultura “Data-driven”: l’instaurazione di una strategia di manutenzione predittiva rappresenta, per il settore industriale, un vantaggio competitivo in quanto permette di sfruttare la raccolta dei dati, grazie a tecnologie di IoT, per la costruzione di una strategia a lungo termine improntata sull’abbattimento dei costi e l’ottimale utilizzo delle risorse disponibili. Ritroviamo questi importanti obiettivi anche nel settore Sanitario dove, a fronte di esigenze sanitarie sempre crescenti (dovute al progressivo invecchiamento della popolazione e al conseguente aumento dei malati cronici) è richiesta una continua evoluzione dei sistemi di cura tale da contenere le spese e rispondere a questi bisogni sanitari ottimizzando l’uso dei mezzi a disposizione. Integrando servizi di manutenzione predittiva, inoltre, si abbracciano quei cambiamenti organizzativi propri della “*Data-driven culture*”, ovvero di tutto l’insieme di approcci, obiettivi, strumenti e *skills* che orbitano intorno al concetto dell’utilizzo ottimale dei dati. Uno degli aspetti chiave di questa modalità di pensiero è proprio l’analisi dei dati generati da, ormai, qualsiasi macchinario, al fine di aumentare la conoscenza di un determinato processo e, in ultima analisi, ottenere su di esso un controllo completo.

Riassumendo quindi: la trasformazione digitale che sta interessando tutti i settori umani (compreso quello sanitario), ovvero la possibilità di raccogliere e conservare

diversi tipi di dati in tempo reale, permette di conoscere meglio i problemi di un macchinario e di evitarne i fermo macchina andando così a migliorare sia, in ambito sanitario, la qualità del servizio offerto dalla macchina in questione sia la gestione della stessa.

3.4 Metodi analitici di predizione

Nei paragrafi precedenti sono stati introdotti obiettivi e vantaggi della manutenzione predittiva. Nel seguente verranno invece descritti i metodi di machine learning sui quali essa si fonda.

Definiamo “Machine Learning” come lo studio o lo sviluppo di modelli e algoritmi che permettono ad un sistema (o un programma) di imparare dall’esperienza per aumentare le proprie performance durante l’esecuzione. Un sistema “impara” dall’esperienza E in riferimento a qualche tipo di task T e ad una misura di performance P, se le sue performance nell’eseguire il task T, come misurate da P, aumentano grazie all’esperienza. La fase di addestramento di un modello viene indicata come fase di *training*. Sistemi di questo tipo vengono chiamati “Modelli Adattativi”.

Un modello è tipicamente una funzione “obiettivo” t_n (funzione “target”) che riceve uno o più (n) input (chiamati anche *features*, *variabili indipendenti*, *osservazioni* o *esempi*) e genera uno o più output (chiamati anche *classi* o *variabili dipendenti*). La definizione di questa funzione viene determinata, nell’approccio proprio del Machine Learning, attraverso un approccio “bottom-up” nel quale il sistema osserva un insieme di esempi e costruisce un modello basato su questi.

A seconda della tipologia di dati usati, si possono identificare tre paradigmi di training:

1. **Addestramento supervisionato** $\langle x_i, t_i \rangle$: per ogni input x_i , dato l’output desiderato t_i , al modello è richiesto di imparare a produrre l’output corretto y dato un input x (mai osservato dal modello in precedenza). Quest’ultimo aspetto determina la capacità di un modello di generalizzare, ovvero produrre output

corretti a fronte di input mai incontrati. Tramite questo tipo di addestramento si cerca di risolvere due tipi di problemi:

- **Classificazione:** Il programma è addestrato per classificare un oggetto in alcune classi (note). Un classico esempio è il problema di identificare quando una mail è classificabile come spam oppure no. Oppure, in ambito sanitario, recenti sono le applicazioni in cui vengono utilizzati algoritmi di machine learning per l'identificazione e la classificazione di una certa patologia in diverse situazioni[9], [10].
 - **Regressione:** Il programma è addestrato per predire un valore numerico e quindi predire l'andamento di una determinata serie numerica. Esempi applicativi di questo tipo possono essere ritrovati in numerosissimi ambiti. Si riportano qui di seguito solo alcuni esempi di applicazioni di machine learning utilizzati in ambito finanziario [11], astronomico [12], turistico [13] e sanitario [14].
2. **Addestramento non supervisionato** $\langle x_i \rangle$: al modello è richiesto di identificare regolarità tra i dati a disposizione per costruire una rappresentazione utilizzabile per condurre analisi e predizioni. Attraverso l'addestramento non supervisionato si possono affrontare problemi di:
 - **Detezione di anomalie:** Il programma analizza i dati a disposizione e impara autonomamente ad identificare pattern anomali.
 - **Clustering:** Il programma identifica gruppi di esempi con caratteristiche simili
 3. **Addestramento rinforzato** $\langle x_i \rangle, \{a_1, a_2, \dots\}, r_i \in R$: Al modello viene richiesto di imparare a produrre azioni a_1, a_2 in modo tale da massimizzare le ricompense ricevute r_i .

Relativamente alla manutenzione predittiva, le applicazioni per cui essa viene applicata rientrano nelle categorie di problemi di Classificazione, di Regressione e di Detezione di anomalie. Nello specifico, a questi problemi può essere associata una domanda che esplicita l'obiettivo che si vuole ottenere tramite la manutenzione predittiva. Le tre domande (d'ora in avanti indicate con il nome di “Use Case”) sono:

1. **Classificazione:** “*La macchina subirà un guasto?*”

2. **Regressione:** “*Tra quanto tempo la macchina subirà un guasto?*”
3. **Detezione di anomalie:** “*Il comportamento della macchina è anomalo?*”

Nei successivi paragrafi si espliciterà ognuno dei seguenti problemi, fornendo alcuni esempi di algoritmi utilizzabili in ognuno dei sopracitati casi.

3.4.1 Classificazione

In questo tipo di problema, come già visto, l’obiettivo è quello di fornire una classificazione per un determinato oggetto. L’“oggetto” in questione si configura come l’eventualità o meno della macchina di subire un guasto, a fronte di una serie di informazioni relative allo stato operativo della macchina stessa. A titolo di esempio, si rimanda al lavoro svolto da Battifarano et al. dal titolo “Predicting Future Machine Failure from Machine State Using Logistic Regression” [15] nel quale, a partire da informazioni riguardanti lo stato operativo della macchina in un dato istante temporale, si è addestrato un modello di regressione logistica per predire lo stato di funzionamento del macchinario con 24 ore di anticipo.

La regressione logistica è un metodo di classificazione comunemente utilizzato in applicazioni simili. Esso modellizza la probabilità di una variabile binaria target \mathbf{Y} (che nel caso sopracitato rappresenta lo stato di “salute” della macchina differenziandolo in “funzionante” e “non funzionante”) a partire da un vettore di features \mathbf{X} (le informazioni sullo stato operativo del macchinario) trasportando una combinazione lineare delle features nell’intervallo (0,1) tramite una trasformazione non lineare data dalla 1:

$$P(Y = 1 \mid X = x) = \pi(x) = \frac{e^{\alpha + \beta^T x}}{1 + e^{\alpha + \beta^T x}} \quad (1)$$

L’addestramento del modello, in questo caso, consiste nell’identificazione del migliore set di parametri α e β , ovvero i valori che aumentano la precisione del modello. Più nel dettaglio, la stima dei coefficienti di regressione della Equation 1 si realizza tramite la risoluzione di un problema di massimizzazione di una specifica funzione, chiamata

“verosomiglianza” L (chiamata anche funzione di *likelyhood* in inglese). Questa funzione esprime la probabilità condizionata che un insieme di esempi (t_1, \dots, t_N) , ovvero i dati utilizzati per l’addestramento del modello, vengano determinati da uno specifico set di parametri θ (in questo caso i parametri sono α e β citati precedentemente).

Prima di inoltrarci nella risoluzione matematica del problema, è importante sottolineare che questo approccio è chiamato frequentista, cioè che fa affidamento esclusivamente sull’esperienza. Una diversa strategia è data, come si vedrà, dall’approccio bayesiano nel quale, per la determinazione dei coefficienti del modello, si utilizza come informazione (mancante nell’approccio frequentista) anche una conoscenza a priori della possibile distribuzione dei parametri ricercati e l’esperienza viene utilizzata di volta in volta per correggere questa conoscenza a priori in modo da allinearsi con le evidenze generate dall’esperienza.

Definiamo quindi il problema di massimizzazione della funzione di likelyhood (chiamato *MLE* ovvero **Maximum Likelihood Estimation**) relativamente ad un problema di classificazione. La funzione obiettivo t_n , nel caso di un problema di classificazione binaria, è una funzione tale che:

$$\begin{aligned} t_n &\in \{0, 1\}, \\ t_n &\sim \text{Be}(y_n) \end{aligned} \tag{2}$$

dalla quale si deduce, per definizione di distribuzione di Bernoulli:

$$p(t_n | x) = y_n^{t_n} (1 - y_n)^{1-t_n} \tag{3}$$

che, a sua volta, permette di definire la funzione di likelyhood:

$$L(\theta) = p(t_1, \dots, t_N | \theta) = \prod_n p(t_n | \theta) = \prod_{n=1}^N y_n^{t_n} (1 - y_n)^{1-t_n} \tag{4}$$

nella quale:

- N rappresenta il numero di osservazioni disponibili
- y_n è la funzione sigmoide relativa all'equazione 1, che può essere riscritta, raccogliendo l'esponenziale al denominatore e sostituendo l'argomento dell'esponenziale, come:

$$\begin{aligned} a_n &= \alpha + \beta^T x, \\ y_n &= \frac{1}{1 + e^{-a_n}} \end{aligned} \tag{5}$$

Di conseguenza, per risolvere il problema di MLE (un problema di massimizzazione) bisogna derivare la funzione di likelyhood 4 rispetto ai parametri e annullarne la derivata così trovata. In riferimento a ciò, si procede applicando prima il logaritmo alla 4 ottenendo:

$$\dot{L}(a_n) = \log(L(\theta)) = \sum_{n=1}^N t_n \log(y_n) + (1 - t_n) \log(1 - y_n) \tag{6}$$

La forma così ottenuta è chiamata log likelyhood. Il problema di massimizzazione di quest'ultima corrisponde al suo duale, ovvero alla minimizzazione della log likelyhood negativa:

$$\begin{aligned} \operatorname{argmax}_{a_n} \dot{L}(a_n) &= \operatorname{argmin}_{a_n} -\dot{L}(a_n) \\ &= \operatorname{argmin}_{a_n} - \left(\sum_{n=1}^N t_n \log(y_n) + (1 - t_n) \log(1 - y_n) \right) \end{aligned} \tag{7}$$

L'ultima espressione è chiamata **cross entropia**. In statistica, la minimizzazione della cross entropia corrisponde alla minimizzazione della divergenza di Kullback-Leibler tra la distribuzione ottenuta e quella target [16] che, in sintesi, rappresenta una misura di quanto due distribuzioni di probabilità siano simili.

La risoluzione della 7 risulta complicata dal punto di vista analitico. Si procede quindi attraverso l'utilizzo di metodi numerici come ad esempio il metodo chiamato "discesa del gradiente". Con questo metodo, si cerca di minimizzare una funzione di costo in

modo iterativo a partire da una soluzione iniziale scelta in modo casuale. Man mano che si procede con le iterazioni k_n , i parametri (indicati nelle prossime equazioni con w) ricercati vengono aggiornati secondo la seguente formula di aggiornamento:

$$w^{k+1} = w^k - \eta \frac{\partial J}{\partial w} \Big|_k \quad (8)$$

dove il secondo termine rappresenta il prodotto tra η , chiamato *learning rate*, e la derivata di una funzione di costo J rispetto ai parametri di interesse, che rappresenta il gradiente della funzione di costo. Nel caso di un problema di classificazione, la funzione di costo è la cross entropia dell'equazione 7. Numericamente, il metodo della discesa del gradiente avviene in questi passaggi:

1. Viene inizializzata una possibile soluzione w^0 in modo casuale
2. Si calcola il gradiente della funzione di costo $\frac{\partial J}{\partial w} \Big|_k$
3. Si aggiorna la soluzione tramite l'equazione di aggiornamento data dall'equazione 8
4. Si ripetono i passaggi 2 e 3 fino alla convergenza.

Tuttavia, il metodo del gradiente presenta delle criticità date dalla forma della funzione di costo e per la risoluzione si rimanda a testi specialistici.

Attualmente, questi metodi di ottimizzazione e di addestramento di modelli predittivi sono implementati in pacchetti e funzionalità di diversi linguaggi di programmazione (Python, R, Matlab) che spesso vengono utilizzati per la costruzione iniziale di questo tipo di applicazioni. Sta quindi all'utilizzatore selezionare il modello più appropriato per il caso in esame.

3.4.2 Regressione

In un problema di regressione, gli output desiderati t_i sono valori continui e l'obiettivo è di predire in modo accurato un nuovo output a partire da nuovi input. Diversamente da quanto accade in un problema di classificazione, quindi, in un problema di regressione si cerca di predire un valore numerico. Un esempio applicativo di manutenzione

predittiva svolta tramite risoluzione di un problema di regressione è dato dal lavoro di Tian et al. [17] nel quale viene utilizzata una rete neurale di tipo feedforward per la stima della vita utile rimanente (in inglese *RUL*, Remaining Useful Life) di alcuni componenti di un macchinario rotante tramite l'analisi delle vibrazioni assorbite dai cuscinetti del macchinario. Le reti neurali sono una tipologia di modello non-lineare caratterizzate da:

- numero di neuroni
- tipologia della rete
- funzione di attivazione
- valori dei pesi sinaptici e dei *bias*

Ritroviamo questi elementi nell'unità funzionale di una rete neurale, ovvero una rete neurale dotata di un singolo neurone, con diversi input e un unico output, chiamata perceptrone. In generale, un neurone artificiale è modellizzato come in figura 3.2 e il suo output dipende dal valore assunto dalla funzione di attivazione, indicata da $g()$.

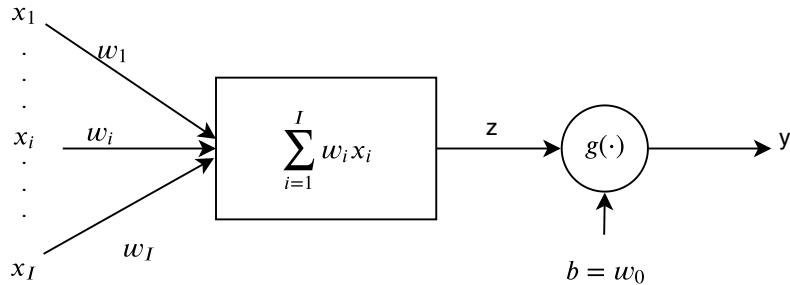


Figura 3.2: Modello di neurone artificiale.

Ispirandosi proprio al neurone biologico e alla sua proprietà di generare un potenziale d'azione secondo la logica del “tutto o nulla”, un neurone artificiale è in grado di replicare l'effetto di sommazione temporale dei potenziali d'azione e l'effetto di “sparo” grazie la definizione di una soglia di attivazione o bias b . L'output di un perceptrone è quindi dato da:

$$y = g \left(\sum_0^I w_i x_i \right) \quad (9)$$

a sua volta, la funzione di attivazione può essere di diversi tipi (gradino, segno, lineare, sigmoide, iperbolica).

L’addestramento di un perceptrone è descritto dal seguente set di equazioni che prende il nome di **Apprendimento Hebbiano**, formulato nel 1949 dallo psicologo canadese Donald Olding Hebb che studiò il meccanismo di apprendimento delle cellule neuronali:

$$\begin{aligned} w_i^{k+1} &= w_i^k + \Delta w_i \\ \Delta w_i &= \eta t x_i \end{aligned} \tag{10}$$

dove:

- η è il learning rate simile a quello dell’equazione 8
- x_i è l’i-esimo input del perceptrone
- t è l’output desiderato

I passaggi dell’apprendimento Hebbiano possono essere sintetizzati nel seguente modo:

1. I pesi vengono inizializzati in modo casuale
2. Si calcola l’output del perceptrone utilizzando i pesi così inizializzati
 2. Se l’output è diverso dall’output desiderato, si applica la regola di aggiornamento data dalla 10 ottenendo un nuovo set di pesi.
3. Si ripete il passaggio 2 fino ad ottenere il risultato corretto.

Tramite questo semplice metodo, un perceptrone è capace di apprendere semplici operazioni logiche quali AND, NOT, OR. Tuttavia, il perceptrone non è in grado di risolvere tutte le operazioni logiche. Infatti mostra limitazioni nell’apprendimento dell’operazione logica XOR. Per ovviare a questo problema, si combinano più neuroni artificiali generando nuovi modelli chiamati Perceptroni Multistrato o, più comunemente, reti neurali feedforward.

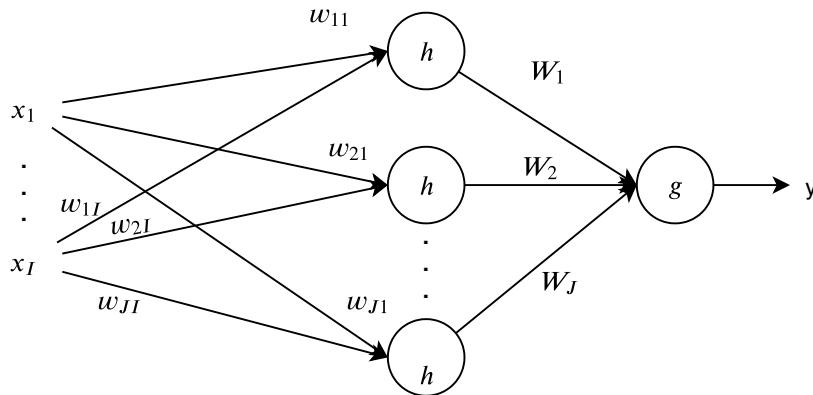


Figura 3.3: Topologia di rete neurale feedforward.

Nelle reti neurali feedforward (figura 3.3), inoltre, sono presenti 1 o più strati nascosti, cosa che non succede nel perceptrone multistrato.

L'output di una rete neurale avente topologia simile a quella in figura 3.3 è:

$$y = g \left(\sum_{j=0}^J W_j h \left(\sum_{i=0}^I w_{ji} x_i \right) \right) \quad (11)$$

Similmente a quanto espresso nel paragrafo precedente riguardante la classificazione, anche in un problema di regressione ci si pone l'obiettivo di minimizzare una funzione di costo (tramite il metodo della discesa del gradiente) che, nel caso di una rete neurale feedforward, è data dalla:

$$E = \sum_{n=0}^N (t_n - y_n)^2 \quad (12)$$

L'apprendimento in una rete neurale feedforward (chiamata così in quanto, per ottenere l'output della rete, l'informazione fluisce unidirezionalmente dallo strato di input a quello di output) avviene tramite un processo chiamato **Backpropagation** che, similmente con quanto visto nel paragrafo precedente, utilizza il metodo del gradiente per minimizzare iterativamente la funzione di errore della rete neurale.

Il processo di backpropagation è costituito da due fasi:

1. La prima fase, chiamata **passo in avanti** consiste nel calcolare, per ogni neurone, l'output dello stesso tramite una funzione non lineare che lega l'input ai pesi associati al neurone considerato.
2. La seconda fase, chiamata **passo indietro** consiste nel ricalcolare tutti i pesi neurali partendo da quelli più vicini allo strato di output fino agli strati di input, utilizzando la regola di aggiornamento data dal metodo del gradiente.

Altre topologie di reti neurali sono le reti neurali ricorrenti e gli autoencoder.

3.4.3 Detezione di anomalie

In questo Use Case si utilizzano dei metodi che sintetizzano i casi di classificazione e regressione visti in precedenza, al fine di costruire dei modelli in grado di monitorare il comportamento di una certa variabile di interesse e stabilire, con una certa precisione, se il comportamento osservato rientra in alcuni gradi di accettabilità definiti a priori.

In estrema sintesi, la detezione di una anomalia può essere effettuata attraverso l'analisi di serie temporali relative ad una grandezza, effettuando delle previsioni sulla serie in esame e classificare quindi il risultato come “Anomalo” o “Normale”. Un esempio di detezione di anomalia tramite analisi di serie temporali è dato dal lavoro svolto da Malhotra et al. [18], nel quale è stata costruita una rete neurale ricorrente in grado di memorizzare delle sequenze numeriche al fine di individuare anomalie nell'andamento di un una certa variabile. Questo tipo di rete neurale viene chiamata Long Short Term Memory (LSTM) per la capacità di, appunto, ricordare sequenze numeriche di lunghezza variabile.

Le reti neurali ricorrenti sono una particolare topologia di rete neurale in cui è inserita una sotto topologia chiamata “rete di contesto” in cui vengono aggiunti nuovi neuroni negli strati nascosti. Gli output di questi neuroni sono connessi sia allo strato di output, sia a agli strati di input implementando un ritardo temporale così come avviene in un sistema retroazionato. L' apprendimento di queste reti neurali

viene svolto attraverso un' estensione del processo standard di backpropagation, chiamato “*backpropagation nel tempo*”. La differenza sostanziale con l'algoritmo standard consiste in una fase preliminare ai passaggi della backpropagation, nella quale la rete ricorrente viene trasformata in una rete neurale feedforward attraverso un procedimento chiamato “*network unfolding*” che procede a “dispiegare” la rete secondo tutti gli istanti temporali dati in input.

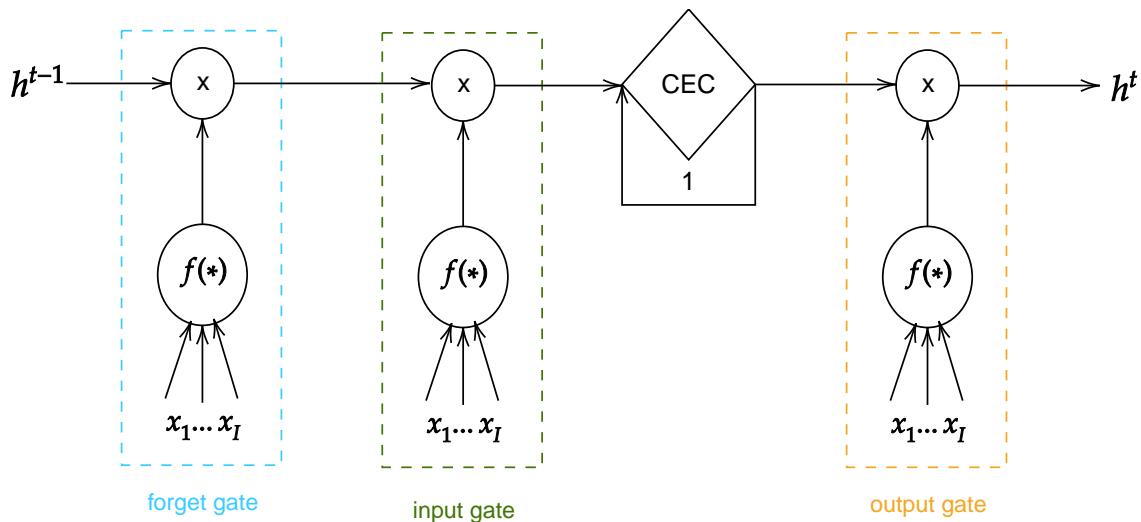


Figura 3.4: Topologia di una LSTM

Una rete di tipo LSTM, come già detto, è una rete neurale ricorrente capace di apprendere sequenze anche molto lunghe. In generale, le LSTM sono composte da sequenze di reti neurali, ognuna di queste caratterizzate da una specifica funzione. La topologia generale di una LSTM viene riportata in figura 3.4.

Si differenziano i seguenti moduli:

- **Forget Gate:** Il risultato di questa rete è 0 o 1. Nel primo caso, quindi, il valore in memoria h^{t-1} viene cancellato. In caso contrario, il valore passa al prossimo modulo in modo inalterato.
- **Input Gate:** Controlla il valore da riportare allo stato h^t .

- **Constant Error Carousel (CEC):** Rete neurale utilizzata per contrastare un problema tipico dell'apprendimento delle reti neurali ricorrenti chiamato “scomparsa del gradiente”. Questo effetto è conseguente all'applicazione del metodo del gradiente durante la fase di network unfolding citata precedentemente. In sintesi, durante lo svolgimento dell'algoritmo dato dalle equazioni 8, si ottiene una catena di moltiplicazioni tra termini tutti inferiori a 1. In questo modo, il gradiente risultante tende ad annullarsi all'aumentare delle fasi di “dispiegamento”, annullando quindi la possibilità di identificare un minimo della funzione di costo. Grazie al CEC, si previene la scomparsa del gradiente perchè si inserisce un elemento che mantiene un peso sinaptico costante e pari a 1.
- **Output Gate:** Ultima rete che controlla il risultato finale della LSTM.

4 La manutenzione predittiva in Ospedale: l'esperienza di Vimercate

In questo capitolo si parlerà dell'applicabilità di una strategia di manutenzione predittiva nell'ambito dell'Ospedale di Vimercate. In questi termini, verrà indicata innanzitutto la modalità di ricerca intrapresa al fine di identificare la classe di dispositivi di maggiore interesse e quella più adatta per essere analizzata con l'obiettivo di progettare e applicare un programma di manutenzione predittiva. Si indagheranno quindi i requisiti necessari per l'applicabilità di questa tipologia di manutenzione, andando poi a valutare il parco macchine presente in ospedale dal punto di vista dell'applicabilità stessa. Nella seconda metà del capitolo verrà quindi riportato il risultato della ricerca preliminare e si evidenzieranno, oltre ai risultati ottenuti dal punto di vista di algoritmi sviluppati, le criticità e le limitazioni incontrate.

4.1 Indagine preliminare

Come visto nel capitolo precedente, la previsione dell'insorgenza di un guasto in un macchinario è ottenibile avendo a disposizione dati ottenuti tramite monitoraggio costante (preferibilmente real time) di variabili fisiche di interesse. Le più comuni variabili sottoposte a monitoraggio, per cui risulta relativamente semplice l'applicazione dello stesso, sono:

- Temperatura: monitorata attraverso termometri dotati di unità di memoria per il salvataggio e la successiva comunicazione in remoto dei dati misurati.
- Pressione: monitorata attraverso sensori applicati alle parti meccaniche di interesse.
- Ampiezza, fase e spettrometria delle vibrazioni meccaniche: variabili, queste, che risultano molto adatte ad essere studiate per il monitoraggio e la previsione dello stato di salute di un dispositivo meccanico sottoposto a movimenti ciclici, o di un componente dello stesso. L'analisi dello spettro delle frequenze si è dimostrato, in letteratura, una buona metodologia di indagine (ref).

- Tensioni e correnti: misurate, manualmente, attraverso amperometri e voltmetri e automaticamente attraverso letture “interne” effettuate dai software stessi dei macchinari.

Per l’individuazione di una classe di dispositivi adeguata per essere trattata e analizzata nell’ottica di manutenzione predittiva, ci si è avvalsi della consultazione sia del Global Service, sia del personale del SIC. I risultati di queste consultazioni possono essere sintetizzati nei seguenti punti:

- L’interesse comune rispetto all’applicazione di una strategia di manutenzione predittiva è indirizzato verso la minimizzazione del fermo macchina in tutte quelle apparecchiature elettromedicali considerati “ad alta incidenza”, ovvero il cui guasto comporta importanti rallentamenti nell’erogazione dei servizi sanitari e l’applicazione clinica per il quale sono utilizzati interessano una grande parte del bacino di utenza della popolazione. Fanno parte di questa famiglia: TAC, Risonanze magnetiche, Mammografi, ecografi, defibrillatori etc. (**esiste una definizione specifica?**)
- Esiste, nel software attualmente utilizzato per la gestione delle apparecchiature elettromedicali (Coswin8i), una funzionalità che permetterebbe di monitorare le prestazioni di alcuni dispositivi ed effettuare quindi una sorta di manutenzione predittiva, ma questa risulta inutilizzata e la documentazione relativa a questa funzionalità risulta del tutto assente.
- Esistono molti dispositivi, in ospedale, che hanno una connessione internet per il collegamento alla rete dell’ospedale, ma nessuno di questi comunica dati fisici utili ad un’analisi di tipo predittivo sullo stato di salute del macchinario. Questa è la maggiore criticità incontrata.
- L’eventuale modifica di un apparecchio elettromedicale, anche se di proprietà dell’ospedale, al fine dell’inserimento di sensori utili alla raccolta di dati fisici (come un semplice termometro) risulterebbe in una invalidazione della certificazione CE, garanzia della sicurezza del dispositivo, innalzando quindi il rischio (sia per il paziente, sia per l’operatore) relativo all’utilizzazione dello stesso.

Si ha quindi una situazione generale dove risulta difficile, per l’Azienda Ospedaliera

in generale e per Ingegneria Clinica in particolare, pensare di poter applicare una metodologia di manutenzione così avanzata. Tuttavia, il sistema informativo presente in Ospedale ha, come si vedrà più avanti nel capitolo 5 relativo allo sviluppo del software , tutti i requisiti per poter, in futuro, ospitare un sistema di manutenzione di questo tipo.

4.2 L'analisi dei log macchina

A fronte della situazione delineata, si è quindi cercato, tramite un processo di ingegneria inversa, di sfruttare sistemi e risorse attualmente disponibili in Azienda, per ottenere dei risultati che, sebbene non siano paragonabili alla “manutenzione predittiva” eseguita in altri settori (come quello manifatturiero), dimostrano l'esistenza di possibilità pratiche, per l'Ospedale, di indirizzarsi su un percorso di profondo sviluppo tecnologico dove la digitalizzazione, la data engineering, l'IoT e applicazioni di intelligenza artificiale hanno e avranno un ruolo predominante.

In questo contesto, è necessario fare riferimento alle variabili fisiche precedentemente descritte e ai sensori responsabili delle loro misurazioni. In realtà, tutte le variabili citate possono essere soggette a monitoraggio automatico “interno” in quei macchinari dotati di un sistema computerizzato. Queste e molte altre letture ed analisi sui processi svolti vengono condotte dai computer, installati all'interno dei macchinari, in modo automatico e a frequenza variabile (impostata dal produttore) per poi essere salvate in file di memoria chiamati “*log*” (o più comunemente log macchina). I log macchina, nei sistemi progettati per produrli, rappresentano un sistema di autodiagnistica utile ai fini del controllo del normale funzionamento del dispositivo da parte del personale tecnico. Si pensi infatti al caso di manutenzione correttiva effettuata sulla risonanza magnetica Philips Achieva (d'ora in avanti chiamata Achieva per brevità) descritto nel capitolo 2.5. In quel frangente, il personale tecnico di Philips ha estratto e caricato i log macchina generati dalla Achieva in un software proprietario in grado di riassumere i log e di fornire al tecnico un rapido riassunto di tutti gli eventi (interni) che hanno interessato la macchina.

La ricerca si è quindi indirizzata, a fronte delle criticità incontrate, nell'analisi dei log macchina al fine di costruire dei modelli predittivi sullo stato di salute del dispositivo che li ha generati.

In particolare, per l'attività si è fatto principalmente riferimento allo studio condotto da Sipos et al. [19] nel quale ci si è avvalsi di diversi Terabytes di log macchina estratti da diverse migliaia di risonanze magnetiche prodotte da Siemens ®. La quantità di dati a disposizione, come si vedrà, rappresenta una dei principali requisiti per la modellizzazione di algoritmi di machine learning perchè, dalla stessa, dipendono fortemente le capacità di apprendimento del modello che si vuole sviluppare.

In questo studio, gli autori evidenziano tre principali problematiche relativo all'utilizzo dei log macchina per la predizione dei guasti. Esse possono essere così riassunte:

1. I log macchina, essendo progettati per supportare personale tecnico nella risoluzione di problemi di tipo informatico (attività di *debugging*), raramente contengono informazioni esplicite per la predizione di un guasto del macchinario.
2. I log contengono dati di tipo eterogeneo tra cui: sequenze simboliche, serie temporali numeriche, testi non strutturati e variabili categoriche. Questa particolarità aggiunge uno strato di complessità all'analisi di questo tipo di dati.
3. I log contengono una grande quantità di dati, il che pone delle sfide dal punto di vista dell'efficienza computazionale.

Elemento imprescindibile per poter effettuare delle analisi predittive è la disponibilità di dati relativi ad interventi di manutenzione passati effettuati sulla macchina in esame. Grazie a questi, infatti, è possibile correlare ogni guasto avvenuto (noto) con il corrispondente insieme di log macchina.

Partendo da questi presupposti si è cercato, in un primo tentativo, di utilizzare i log macchina derivanti dalla Achieva estratti in concomitanza con l'intervento di manutenzione correttiva descritto nel capitolo (REF). Tuttavia, i dati a disposizione estratti facevano riferimento a pochi giorni di attività della macchina. Non sufficienti, quindi, ad essere utilizzati per condurre un'analisi che desse buoni risultati. Si è presentata quindi la necessità di reperire dalla macchina stessa altri log. Tuttavia

sono state incontrate diverse difficoltà legate alla modalità di ottenimento di tali dati. Sinteticamente, per accedere ai log macchina è necessario, infatti, accedere dal terminale della risonanza magnetica tramite le credenziali di accesso in dotazione esclusiva ai tecnici e ai field engineers autorizzati (di Philips).

4.3 Le lavaendoscopi MEDIVATORS ®ISA ®

A seguito di tali limitazioni, si è quindi spostata l'attenzione su un sistema più semplice rispetto ad una risonanza magnetica, ma caratterizzata da una modalità di generazione e salvataggio di log macchina utili ai fini predittivi. Su suggerimento di uno degli ingegneri biomedici del SIC, è stata individuata una famiglia di dispositivi medici chiamate lavaendoscopi, su cui poi è stato sviluppato un software di monitoraggio comprensivo di un modulo di manutenzione predittiva. In figura 4.1 è riportata una fotografia del dispositivo in questione.



Figura 4.1: Lava-Sterilizzatrice MEDIVATORS®ISA®

Nello specifico, la Lava-Sterilizzatrice MEDIVATORS® ISA® è un dispositivo medico progettato per il lavaggio e la sterilizzazione chimica a freddo degli endoscopi rigidi e flessibili e degli accessori endoscopici.

In sintesi, il processo di utilizzo della macchina consiste in diverse fasi:

1. Accensione della macchina
2. Carico endoscopi nella vasca
3. Chiusura della vasca e selezione del ciclo di riprocessazione desiderato

4. Prelievo dell'endoscopio dalla vasca

In concomitanza con la fine di un ciclo di lavaggio, il dispositivo registra nel proprio hard disk tutte le informazioni relative a tutti ai cicli eseguiti creando un archivio elettronico consultabile in qualsiasi momento. Inoltre, è dotato di una stampante integrata che, al termine di ogni ciclo, stampa in automatico un report del ciclo. Il report è un documento essenziale per la convalida del ciclo e deve essere sempre conservato.

L'attenzione è stata posta proprio su questi report di stampa chiamati per brevità "scontrini". In figura 4.2 viene riportato un esempio di scontrino in formato digitale.

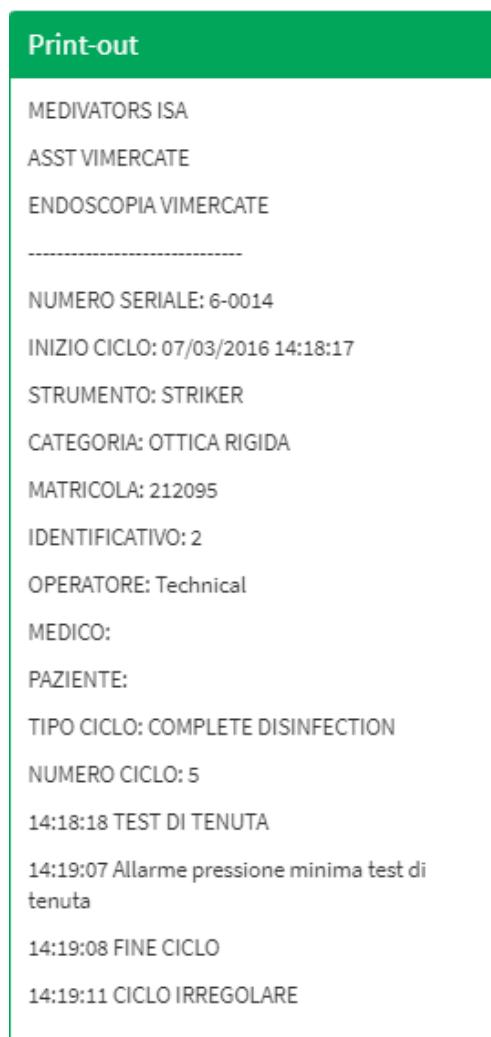


Figura 4.2: Report stampato dalla lava-sterilizzatrice MEDIVATORS ®ISA ®

Sempre in riferimento alla figura 4.2, i parametri inseriti in stampa sono:

- Data ed ora di inizio ciclo
- Dati strumento (categoria-s/n)
- Medico (opzionale)
- Paziente (opzionale)
- Tipo di ciclo eseguito

- Numero progressivo del ciclo
- Fasi del ciclo con relativi tempi di contatto e temperatura
- Esito del ciclo

Gli scontrini, quindi, rappresentano una buona fonte di dati per quanto riguarda lo stato di funzionamento della macchina in quanto, come visto, essi riportano sia gli attori coinvolti nello specifico ciclo di lavaggio, sia gli eventuali allarmi registrati durante il lavaggio. Essi riportano inoltre variabili numeriche quali temperatura e tempi delle varie fasi del ciclo selezionato.

Con questi dati a disposizione si è indagata quindi la possibilità di prevedere, con un anticipo di 7 giorni, l'insorgenza di guasti tali da indurre il personale del reparto a richiedere un intervento di manutenzione correttiva al Global Service.

L'attività svolta si è articolata in diverse fasi, descritte nel dettaglio nei successivi paragrafi. Lo schema delle operazioni eseguite è riassunto nella figura 4.3.



Figura 4.3: Schema di lavoro

4.4 Raccolta Dati

La prima fase operativa è stata quella di raccolta ed estrazione degli scontrini dalle macchine in questione. Grazie alla responsabile del reparto di endoscopia e ad uno dei collaboratori tecnici del SIC, è stato possibile estrarre da una MEDIVATORS®ISA®, l'intero storico dei report di lavaggio conservati nell'hard disk della macchina per un totale di 5441 scontrini (pari a 3 anni di attività). I dati estratti sono stati salvati su una chiavetta USB e l'estrazione ha impiegato circa 20 minuti.

4.5 Conversione dei file di backup

Per utilizzare le informazioni contenute negli scontrini estratti dalla memoria della macchina, sono stati scritti ed utilizzati diversi script utilizzando il linguaggio di programmazione R. Il primo tra questi ad essere progettato è stato quello responsabile della trasformazione dei 5441 file di testo in un unico file tabulare attraverso diverse funzioni scritte ad hoc. Questo script, chiamato “Analyzer.R” individua gli scontrini e riorganizza le informazioni estraibili da questi in un formato a righe e colonne (questo tipo di dato viene chiamato “dataframe” in R). In particolare, ogni riga corrisponde ad uno scontrino e ogni colonna rappresenta una “feature” identificabile nello scontrino stesso. Questo passaggio è stato necessario al fine di disporre di una struttura dati ben organizzata e coerente, indispensabile per le successive fasi di modellizzazione. La vera e propria conversione del dato avviene per mezzo di particolari pattern di ricerca chiamate “espressioni regolari” o “Regex” (Regular expression). Si è scelto di usare questo tipo di espressioni in quanto, sebbene tutti e i 5441 scontrini fossero file di testo diversi l’uno dall’altro, si basano sul concetto di sfruttare delle regolarità presenti in un dato testo ed estrarne informazioni di interesse. In sintesi, per la conversione degli scontrini si è organizzata la struttura degli stessi in 3 parti, per le quali sono state scritte funzioni di estrazione utilizzando le Regex citate precedentemente. La struttura degli scontrini è stata così separata:

1. **Intro:** parte introduttiva dello scontrino che reca il nome della macchina (Medivators ISA), il nome della ASST e il reparto di ubicazione della macchina. Queste righe, comuni a tutti gli scontrini, non apportano contenuto informativo utile ai fini predittivi e sono quindi stati scartati.
2. **Header:** questa parte di testo contiene alcune delle informazioni precedentemente elencate in riferimento alla figura 4.2. Queste informazioni sono presentate (in tutti gli scontrini) con la stessa “struttura” ovvero: NOME CAMPO: VALORE CAMPO. Di conseguenza sono state utilizzate due espressioni regolari per “estrarre” rispettivamente il NOME e il VALORE del campo:
 - NOME CAMPO: per estrarre il nome del campo si è utilizzata la seguente

regex: `.*(?=:)`. I simboli utilizzati in una regex vengono chiamati “metacaratteri” e, nel caso della regex precedente, essi possono essere tradotti con la seguente istruzione: “Seleziona tutti i caratteri (tramite i metacaratteri `.*`) che precedono un carattere di due punti e spazio (tramite la metaistruzione chiamata ‘positive lookahead’ `(?=:)`).”

- **VALORE CAMPO:** per estrarre invece il valore del campo si è utilizzata la seguente regex: `(?<=:).*` che è esattamente speculare alla prima. In questa infatti vengono selezionati i caratteri che seguono il carattere “`:`” e lo spazio.
- 3. **Footer:** in questa parte sono elencate le varie fasi di lavaggio il cui nome è anteposto ad un timestamp corrispondente all’orario di inizio della fase stessa. Anche in questo caso, per estrarre il nome delle fasi di lavaggio, sono state utilizzate delle regex in modo tale da “estrarre” i caratteri che si trovavano in seguito ad un timestamp. In particolare, si è utilizzata la regex: `\d\d:\d\d:\d\d.*`

In figura 4.4 sono riportati due scontrini ed evidenziate le tre parti descritte.

MEDIVATORS ISA ASST VIMERCATE ENDOSCOPIA VIMERCATE	Intro	MEDIVATORS ISA ASST VIMERCATE ENDOSCOPIA VIMERCATE
NUMERO SERIALE: 6-0014 INIZIO CICLO: 05/10/2016 15:33:49 STRUMENTO: VNL-1170 K CATEGORIA: LARINGOSCOPIO MATRICOLA: G120201 IDENTIFICATIVO: 21 OPERATORE: G.B. MEDICO: PAZIENTE: TIPO CICLO: STERILIZZAZIONE VELOCE NUMERO CICLO: 1071	Header	NUMERO SERIALE: 6-0014 INIZIO CICLO: 25/10/2016 08:01:13 STRUMENTO: 0 CATEGORIA: 0 MATRICOLA: IDENTIFICATIVO: 0 OPERATORE: C.G. MEDICO: PAZIENTE: TIPO CICLO: AUTOSANIFICAZIONE NUMERO CICLO: 1170
15:33:49 TEST DI TENUTA 15:34:49 CARICO ACQUA 15:36:07 TEMPERATURA:24°C 15:36:28 CARICO STERILIZZANTE 1 15:37:19 CAMBIO TANICA 15:42:59 CARICO STERILIZZANTE 2 15:43:16 CARICO STERILIZZANTE 2 15:44:03 STERILIZZAZIONE DINAMICA 15:54:11 SCARICO 15:55:50 CARICO ACQUA 15:56:30 RISCIACQUO 15:58:18 SCARICO 15:59:56 ASCIUGATURA 16:00:43 FINE CICLO 16:04:08 PRELEVATO DA: M.F. 16:04:08 CICLO REGOLARE	Footer	08:01:13 TEST DI TENUTA 08:01:33 CARICO ACQUA 08:02:35 CARICO STERILIZZANTE 1 08:02:42 TEMPERATURA:29°C 08:03:26 CARICO STERILIZZANTE 2 08:04:12 MISCELAZIONE 08:05:20 DISINFEZIONE 08:15:22 SCARICO 08:16:51 CARICO ACQUA 08:17:23 RISCIACQUO 08:19:03 SCARICO 08:20:32 ASCIUGATURA 08:21:48 FINE CICLO 08:24:32 PRELEVATO DA: B.M. 08:24:32 CICLO REGOLARE

Figura 4.4: Divisione degli scontrini nelle parti di Intro, Header e Footer

Con operazioni simili a quelle descritte, ovvero analizzando il testo degli scontrini per trovare e sfruttare le regolarità presenti nel testo, si è costruita una funzione che ricevesse come input un file in formato .txt, rappresentante lo scontrino, e che restituisse in uscita un dataframe composto da una sola riga. La funzione è stata poi fatta ciclare su tutto l'insieme di scontrini a disposizione ottenendo un'unica tabella da 5441 righe e 12 colonne (features). Un esempio del dataframe generabile dalla funzione descritta è riportato in figura 4.5.

INIZIO CICLO	TIPO CICLO	NUMERO SERIALE LAVENDOSCOPI	MODELLO DELLO STRUMENTO	CATEGORIA	NUMERO SERIALE STRUMENTO	IDENTIFICATIVO	OPERATORE	NUMERO CICLO	ALLARMI	ESITO. CICLO
2019-03-11 09:33:40	STERILIZZAZIONE COMPLETA	6-0540	GIF-Q180	Gastroscope	2800071	01	E.V.I.	9	Nessun allarme rilevato	CICLO REGOLARE
2019-03-11 09:39:36	STERILIZZAZIONE COMPLETA	6-0539	CF-Q165I	Colonoscope	2802049	11	E.V.I.	10	Nessun allarme rilevato	CICLO REGOLARE
2019-03-11 09:52:23	AUTOSANIFICAZIONE	6-0538	0	0	0	T294	2	Nessun allarme rilevato	CICLO IRREGOLARE	
2019-03-11 09:52:55	AUTOSANIFICAZIONE	6-0538	0	0	0	T294	3	Nessun allarme rilevato	CICLO REGOLARE	
2019-03-11 09:56:36	STERILIZZAZIONE COMPLETA	6-0537	EB-1970K	Bronchoscope	G120511	24 - B	I.O.	13	Nessun allarme rilevato	CICLO REGOLARE
2019-03-11 10:21:08	STERILIZZAZIONE COMPLETA	6-0540	CF-Q165I	Colonoscope	2802043	12	N.R.	10	Nessun allarme rilevato	CICLO REGOLARE

Figura 4.5: Tabella ottenuta dalla conversione dei file testuali degli scontrini.

È importante sottolineare il fatto che, tra le feature selezionate, ne è stata creata una che contenesse il testo dello scontrino “pulito” dai caratteri di punteggiatura e caratteri non alfanumerici. Il motivo di questa scelta sta nella decisione di sfruttare la natura “testuale” di questi log macchina utilizzando tecniche di text mining per costruire dei modelli predittivi basandosi esclusivamente sulle parole contenute nei diversi file testuali, andando a sfruttare quindi la struttura originaria del log macchina.

4.6 Modellizzazione

La fase di modellizzazione è stata sicuramente quella in cui sono state incontrate più difficoltà, principalmente dovute, come si vedrà, ad un non sufficiente numero di dati a disposizione. Innanzitutto, è stato necessario estrarre dal software di manutenzione “Coswin8” utilizzato in ospedale, lo storico delle manutenzioni effettuate sulla macchina da cui sono stati estratti gli scontrini. Avendo a disposizione questi ultimi dati, si sono potuti incrociare le date presenti sugli scontrini con quelle riportate dal servizio di manutenzione, così quindi di avere la possibilità di studiare gli scontrini corrispondenti a 7 giorni precedenti all’effettiva chiamata al Global Service (d’ora in avanti chiamati

per brevità “giorni predittivi”). Si è cercato quindi, nel seguente lavoro, di correlare i guasti della lavaendoscopi alle informazioni contenute negli scontrini dei giorni predittivi, al fine di ottenere un modello di predizione che calcolasse la probabilità di guasto della lavaendoscopi sulla base dello storico dato dai 3 anni di informazioni a disposizione.

4.6.1 Apprendimento ad istanza multipla

Per costruire tale modello è stata quindi indagato l’approccio di apprendimento migliore, sulla base di quanto descritto nel capitolo 3.4, ed è stato scelto, sulla base di un precedente lavoro trovato in letteratura , una tipologia di apprendimento supervisionato chiamato “Apprendimento ad istanza multipla” (Multiple instance learning, MIL). In questa metodologia, il classificatore non riceve una serie di esempi indipendenti x_i caratterizzati da una etichetta (l’output desiderato t_i) come avviene nel classico apprendimento supervisionato, ma riceve un insieme di “borse” o “contenitori” a cui viene associata un’etichetta che rappresenta l’output che dovrà imparare a produrre. Ogni borsa può contenere più istanze (che in questo caso sono gli scontrini di un giorno di attività) e, nello specifico, una borsa viene etichettata come “positiva” se contiene al suo interno scontrini appartenenti ad un “giorno predittivo”, mentre invece viene etichettata come “negativa” altrimenti. Di conseguenza, il classificatore desiderato è stato progettato per classificare una borsa di scontrini come positiva o negativa, a seconda delle informazioni contenute negli scontrini di un singolo giorno. Uno schema riassuntivo dell’approccio descritto è apprezzabile in figura 4.6.

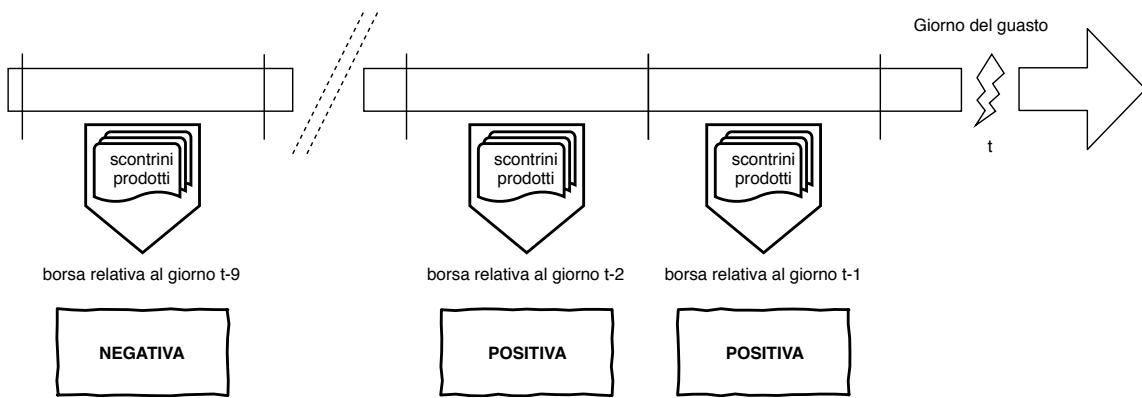


Figura 4.6: Apprendimento ad istanza multipla. La prima borsa è associata ad una etichetta “negativa” in quanto al di fuori dell’intervallo di predizione dei 7 giorni.

Successivamente al raggruppamento degli scontrini in borse, si è cercato di risolvere questo problema di classificazione riducendo la complessità generale passando da un approccio MIL a quello più standard di apprendimento supervisionato. Questo è stato reso possibile grazie alla trasformazione di tutti gli scontrini presenti in una singola borsa in un unico “metascontrino” (il cui nome specifico è quello di “metaesempio”), ovvero una osservazione unica generata tramite la sintetizzazione delle informazioni presenti in tutti gli scontrini della stessa borsa. Così facendo, quindi, non si hanno più degli insiemi di scontrini contenenti diverse osservazioni, ma, allineandosi quindi al più tradizionale approccio di apprendimento supervisionato, una serie di osservazioni (i metascontrini) caratterizzate da una etichetta. Quest’ultima struttura risultante (associabile ad un dataframe di righe e colonne) risulta quindi più facile da gestire per quanto riguarda la modellizzazione di un algoritmo di predizione.

4.6.2 Text mining: creazione del Corpus

Per creare questi “metascontrini” sono state utilizzate diverse tecniche di text mining sfruttando le capacità della libreria R chiamata “tm”, diminutivo di “text mining”. A partire dalla feature chiamata “testo”, creata dalla funzione descritta nel paragrafo 4.5,

è stato infatti creato un “corpus” ovvero una collezione di documenti. In particolare, convertendo tutti gli scontrini presenti, il corpus ottenuto è un oggetto composto da 5441 documenti, ognuno dei quali composto da diversi caratteri. Esplorando i primi 5 documenti del corpus, per esempio, si ottiene:

```
<<VCorpus>>
Metadata: corpus specific: 0, document level (indexed): 0
Content: documents: 5

[[1]]
<<PlainTextDocument>>
Metadata: 7
Content: chars: 115

[[2]]
<<PlainTextDocument>>
Metadata: 7
Content: chars: 90

[[3]]
<<PlainTextDocument>>
Metadata: 7
Content: chars: 55

[[4]]
<<PlainTextDocument>>
Metadata: 7
Content: chars: 290

[[5]]
<<PlainTextDocument>>
Metadata: 7
Content: chars: 115
```

Nell'output precedente è intuitibile la struttura del corpus. Ogni elemento dello stesso viene identificato come un “PlainTextDocument” il cui contenuto è dato da un variabile numero di caratteri testuali. Esplorando ad esempio il primo documento è possibile leggere il testo del relativo scontrino, pulito precedentemente da eventuali caratteri speciali, punteggiatura e numeri:

```
<<PlainTextDocument>>
Metadata: 7
Content: chars: 115

tipo ciclo complete disinfection test di tenuta allarme pressione minima test di
tenuta fine ciclo ciclo irregolare
```

4.6.3 Text mining: Document Term Matrix

Il passo successivo è stato quello di convertire il corpus creato in una struttura a matrice tipicamente utilizzata in applicazioni di text mining chiamata “Document Term Matrix”. Questo tipo di matrice è composta da un numero di righe pari al numero di documenti presenti nel corpus e da tante colonne quanti sono i termini presenti in tutti i documenti, presi una sola volta. Il valore all'incrocio di una riga e di una colonna determina la tipologia di “peso” che viene assegnato ad ogni termine rispetto ad ogni documento. Tra le “pesature” possibili, si riportano quelle più comunemente utilizzate in questo tipo di applicazioni:

1. **Term Frequency (tf)**: indicata con $tf(t, d) = f_{t,d}$ rappresenta il conteggio delle volte in cui è presente un termine t in un documento d .
2. **Term Frequency - Inverse Document Frequency (tf-idf)**: definita come il prodotto tra la Term Frequency ed la Inverse Document Frequency. Quest'ultima è una misura del quantitativo informativo che una parola apporta rispetto alla sua frequenza in un set di documenti. Matematicamente è definita come:

$$\text{idf}(t, D) = \log\left(\frac{N}{n_t}\right) \quad (13)$$

dove:

- N rappresenta il numero di documenti nel corpus
- n_t è il numero di documenti in cui compare il termine t

Di conseguenza, la tf-idf è definita come:

$$\text{tfidf}(t, d, D) = \text{tf}(t, d) \cdot \text{idf}(t, D) = f_{t,d} \cdot \log\left(\frac{N}{n_t}\right) \quad (14)$$

In figura 4.7 viene riportato un esempio di come la pesatura tramite tf-idf tende a configurarsi come un “filtro” per i termini più comuni, prendendo come esempio due termini presenti in uno stesso documento con la stessa frequenza. Si osserva, infatti, che a pari frequenza corrisponde un punteggio che è maggiore per i termini che compaiono meno frequentemente nel corpus.

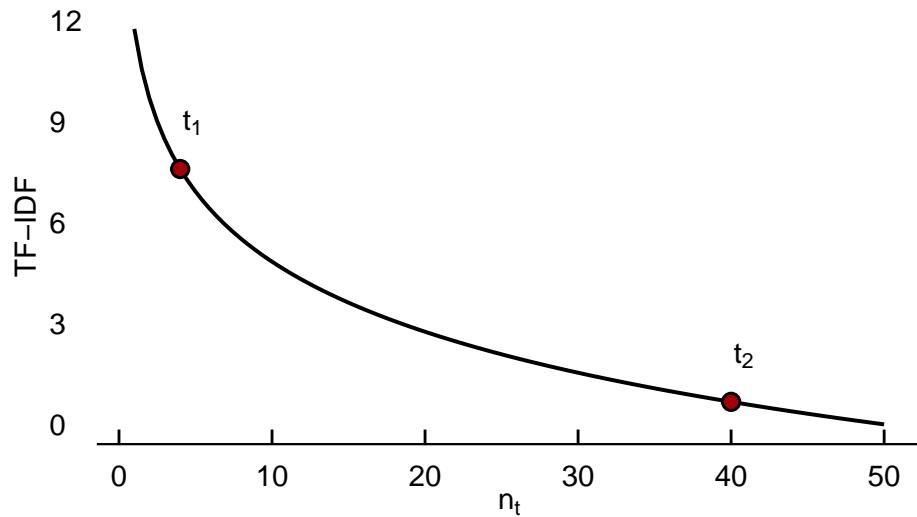


Figura 4.7: Valore dato dalla tf-idf per due termini generici t_1 e t_2 , ipotizzando la stessa frequenza ($f_{d,t_{1,2}} = 3$) all'interno di un generico documento d . t_1 compare in meno documenti del corpus ($n_{t_1} = 4$), di conseguenza il suo peso risulta maggiore.

Tra le due metodologie di pesatura elencate, è stata scelta la seconda. Questo perché la natura degli scontrini può essere meglio caratterizzata dai termini relativi

a guasti e allarmi, che sono infatti meno frequenti rispetto a quelli che indicano un comportamento nominale dell'apparecchiatura. La frequenza dei termini “tipici” di uno scontrino indicante un comportamento anomalo è evidenziata nel grafico a barre in figura 4.8.

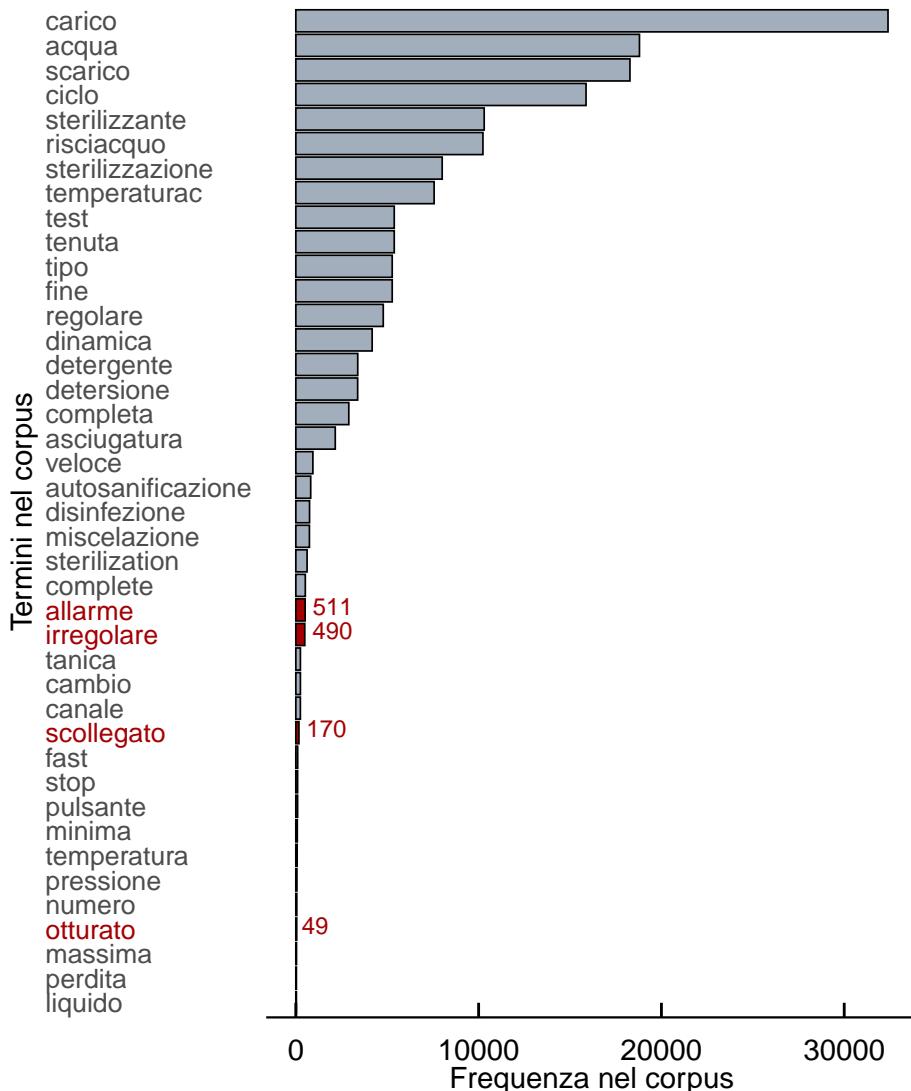


Figura 4.8: Frequenza dei 40 termini più comuni nel corpus degli scontrini.

La rarità dei termini che riflettono un comportamento anomalo è diretta conseguenza

della poca frequenza con cui si verificano eventi di guasto. Dal software di gestione della manutenzione, infatti, è stato riscontrato che la macchina in questione è stata interessata da 47 interventi di manutenzione correttiva. Su circa 800 giorni di attività della macchina, quindi, i guasti sono da considerarsi come eventi estremamente rari, interessando infatti poco più del 5% dei giorni di attività totali.

È parso ragionevole, quindi, al fine di discriminare in modo ottimale un comportamento anomalo da uno nominale, assegnare un peso maggiore ai termini con una frequenza relativa minore tramite la tf-idf descritta sopra.

A seguito di queste considerazioni, si riportano, a solo scopo di esempio, le prime 5 righe di una Document Term Matrix ottenuta dalla trasformazione del corpus.

Tabella 4.1: Document Term Matrix

	allarme	ciclo	disinfezione	irregolare	regolare	sterilizzante
Scontrino 1	0.2403061	0	0.0000000	0.229516	0.0000000	0.0000000
Scontrino 2	0.0000000	0	0.0000000	0.286895	0.0000000	0.0000000
Scontrino 3	0.0000000	0	0.0000000	0.0000000	0.0000000	0.0000000
Scontrino 4	0.0000000	0	0.0831385	0.0000000	0.0045699	0.0059028
Scontrino 5	0.2403061	0	0.0000000	0.229516	0.0000000	0.0000000

4.6.4 Fase di training e testing

Una volta raggruppati gli scontrini in borse corrispondenti a 7 giorni di attività, ad ogni borsa è stata assegnata la label “positiva”, se tra gli scontrini della borsa era presente almeno uno corrispondente ad uno dei giorni “predittivi”, definiti nel paragrafo 4.6.1, “negativa” altrimenti. Le borse sono poi state ridotte ad un'unica osservazione (metascontrino) mediando i valori della DTM per ogni termine presente al suo interno.

Successivamente, i dati sono stati divisi in gruppo di training e testing (chiamati d'ora in avanti training e test set), sui quali è stato costruito il modello di predizione. Per il processo di training è stato utilizzato un pacchetto di R chiamato “caret” (acronimo

di “Classification And REgression Training”), contenente un gruppo di funzioni che facilitano la creazione di modelli predittivi tramite strumenti per la divisione dei dati in gruppo di training e testing (data splitting), pre-processamento dei dati, ottimizzazione dei modelli (model tuning) e altre funzionalità.

Specificatamente, nella fase di training, le funzioni nel pacchetto caret si occupano di:

- valutare, ripetendo le operazioni di addestramento di un modello tramite riselezione dei dati (resampling), l'effetto dei parametri di ottimizzazione (ovvero i parametri specifici di un modello, come α e β della di una regressione logistica, descritti nel capitolo 3.4.1) sulle performance di predizione
- scegliere il modello “ottimale” tra questi parametri
- stimare le performance del modello dal training set

In particolare, per agevolare la fase di training e test, è stata creata una funzione chiamata “model_maker” in grado di allenare diversi modelli, definiti da una lista scelta dall’utente, in modo sequenziale e di salvare le performance di predizione di ogni modello in una tabella consultabile a fine addestramento. Questo ha aiutato nella scelta del modello ottimale, in quanto è stato possibile confrontare diversi risultati di addestramento al variare di alcune opzioni di modellizzazione come ad esempio il metodo di resampling, la scelta dei modelli e la modalità di creazione dei metascontrini.

4.7 Scelta del modello ottimale

Prima di introdurre i risultati, si riportano di seguito le definizioni delle statistiche utilizzate per la valutazione delle performance dei modelli sviluppati. Si utilizza come riferimento una matrice di confusione, struttura spesso utilizzata per valutare le performance di un modello di classificazione binaria (anche se facilmente estendibile al caso di classificazione multipla) caratterizzata dalla seguente struttura:

		Tabella 4.2: Matrice di confusione	
		Riferimento	
		Negativi	Positivi
Previsione	Negativi	TN ¹	FP ³
	Positivi	FN ²	TP ⁴

¹ True Negative: osservazioni classificate come negative, in modo corretto

² False Negative: osservazioni classificate come negative, erroneamente

³ False Positive: osservazioni classificate come positive, in modo corretto

⁴ True Positive: osservazioni classificate come positive, erroneamente

Dalla tabella è possibile definire le seguenti statistiche:

-**accuratezza**: rapporto tra le osservazioni classificate correttamente e il numero totale di osservazioni

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

-**sensitività**: chiamata anche recall o true positive rate, è definita come il rapporto tra i veri positivi e i positivi totali

$$Sensitivity = \frac{TP}{TP + FN}$$

-**specificità**: definita come il rapporto tra i veri negativi e i negativi totali

$$Specificity = \frac{TN}{TN + FP}$$

-precisione: definita come il rapporto tra i veri positivi e la somma tra veri positivi e falsi positivi

$$Precision = \frac{TP}{TP + FP}$$

In un caso come questo di manutenzione predittiva, in cui gli eventi di guasto rappresentano la classe caratterizzata da una frequenza minore rispetto al normale comportamento del macchinario, la statistica di accuratezza non è un buon indice per valutare la bontà di un classificatore che ha come compito quello di identificare, appunto, comportamenti anomali. In casi come questi, in cui è presente uno sbilanciamento tra classi, è meglio fare affidamento su altre statistiche come sensitività, specificità e precisione, perché esse tengono meglio conto della capacità del classificatore di discriminare tra eventi “positivi”, di guasto, e “negativi”.

Di seguito si riportano le performance di predizione dei modelli ottenuti.

Tabella 4.3: Performance dei modelli ottenuti

Model name	Accuracy	Sensitivity	Specificity	Precision	MCC
Bayesian Generalized Linear Model	0.75	0.23	0.92	0.47	0.19
Generalized Linear Model	0.39	0.71	0.29	0.25	0.00
Naive Bayes	0.73	0.11	0.93	0.36	0.08
Neural Network	0.71	0.11	0.91	0.29	0.03
Support Vector Machine (Linear Kernel)	0.73	0.20	0.91	0.41	0.14

Sebbene l'accuratezza risulti buona, essa, come già detto, è inaffidabile in un caso come questo, dove la distribuzione delle classi da predire risulta “sbilanciata”. Tutti i modelli addestrati si dimostrano buoni classificatori per quanto riguarda l'identificazione della classe “negativa”. Tuttavia, come intuibile dal valore di sensitività molto basso caratteristico a tutti i modelli, essi non risultano essere in grado di discriminare la classe di interesse. Ciò è dovuto molto probabilmente al numero eccessivamente basso di osservazioni a disposizione per l'addestramento dei modelli.

Di per sé , infatti, lo sbilanciamento delle classi è un problema facilmente aggirabile

tramite diversi approcci come ad esempio specificando al pacchetto caret di utilizzare un metodo di resampling che tenga conto del forte sbilanciamento di classi (downsampling). In questo modo, ad esempio, ad ogni iterazione vengono selezionate un pari numero di osservazioni positive e negative (utilizzando tutte quelle positive a disposizione), annullando quindi lo sbilanciamento.

In questo lavoro, tuttavia, accorgimenti di questo tipo non hanno migliorato in modo sufficiente le performance di classificazione, ottenendo comunque sensitività e specificità molto basse. Performance migliori possono essere ottenute, banalmente, aumentando il numero di dati a disposizione prelevando da altre macchine simili i file di backup. Elemento essenziale per poter meglio sviluppare questi modelli sarà, comunque, la presenza di uno storico dato da un software di gestione delle manutenzioni senza il quale, infatti, non si avrebbe la possibilità di etichettare i dati a disposizione, invalidando quindi il concetto di apprendimento supervisionato.

5 INSIGHT: IoT al servizio del reparto di endoscopia

In questo capitolo viene descritto in che modo si è arrivati alla decisione di sviluppare una web app (chiamata INSIGHT) per il reparto di endoscopia. Vengono prima indicate le problematiche relative alla riprocessazione della strumentazione endoscopica e, per ognuna di esse, verrà descritto in che modo la web app sviluppata può essere d'aiuto nella razionalizzazione e nella successiva risoluzione delle stesse. Verrà inoltre trattata la strategia utilizzata per effettuare il *software deployment* che ha reso disponibile l'applicazione sulla rete interna dell'ospedale, grazie al contributo dell'U.O.C. Sistemi Informativi Aziendali. Infine si evidenzieranno sia le criticità incontrate durante lo sviluppo del software, sia il riscontro all'utilizzo dell'app da parte del personale del reparto.

6 Conclusioni

In questo capitolo finale verranno riassunti i risultati ottenuti, a fronte delle criticità e dei problemi incontrati durante tutte le fasi dell’esperienza di tirocinio. Successivamente, il capitolo tratterà degli sviluppi futuri riguardo sia la parte di manutenzione predittiva studiata, sia quella legata al software sviluppato. L’ultima parte del capitolo sarà invece dedicata ai ringraziamenti e ad alcune riflessioni personali sull’esperienza maturata in seguito allo svolgimento dei mesi di tirocinio.

A fronte della situazione delineata, si è quindi cercato, tramite un processo di ingegneria inversa, di sfruttare sistemi e risorse attualmente disponibili in Azienda, per ottenere dei risultati che, sebbene non siano paragonabili alla “manutenzione predittiva” eseguita in altri settori (come quello manifatturiero), dimostrano l’esistenza di possibilità pratiche, per l’Ospedale, di indirizzarsi su un (ormai inevitabile) percorso di profondo sviluppo tecnologico dove la digitalizzazione, la data engineering, l’IoT e applicazioni di intelligenza artificiale hanno e avranno un ruolo predominante.

Bibliografia

- [1] “Azienda Socio-Sanitaria Territoriale (ASST) di Vimercate.” <http://www.asst-vimercate.it/web/index.php/organizzazione/organigramma>.
- [2] P. Costruzioni, “Pessina costruzioni 2009 - ospedale vimercate monografia.”
- [3] “Gazzetta Ufficiale.”
- [4] I. C. ASST Vimercate, “Procedura manutenzione programmata.” 2013.
- [5] I. C. ASST Vimercate, “Procedura manutenzione correttiva.” UOC Ingegneria Clinica ASST Vimercate, 2013.
- [6] B. Schmidt, U. Sandberg, and L. Wang, “NEXT GENERATION CONDITION BASED PREDICTIVE MAINTENANCE,” p. 8.
- [7] M. Maccarelli, “Manutenzione: tutte le definizioni delle norme di riferimento,” *Certifico Srl.* <https://www.certifico.com/normazione/173-documenti-riservati-normazione/documenti-estratti-norme/3135-manutenzione-tutte-le-definizioni-delle-norme-di-riferimento>.
- [8] “Prestazioni Ambulatoriali | Open Data Regione Lombardia.” <http://tiny.cc/gfa88y>.
- [9] M. S. Manak *et al.*, “Live-cell phenotypic-biomarker microfluidic assay for the risk stratification of cancer patients via machine learning,” *Nature Biomedical Engineering*, vol. 2, no. 10, p. 761, Oct. 2018.
- [10] P. Wang *et al.*, “Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy,” *Nature Biomedical Engineering*, vol. 2, no. 10, p. 741, Oct. 2018.
- [11] “Stock performance modeling using neural networks: A comparative study with regression models - Semantic Scholar.” <https://www.semanticscholar.org/paper/Stock-performance-modeling-using-neural-networks%3A-A-Refenes-Zapranis/17ec85347dd8950d2895b6b6ed6fee3ac>

- [12] T. Nagem, R. Qahwaji, S. Ipson, Z. Wang, and A. Al-Waisy, “Deep Learning Technology for Predicting Solar Flares from (Geostationary Operational Environmental Satellite) Data,” *International Journal of Advanced Computer Science and Applications*, vol. 9, Feb. 2018.
- [13] S. Sun, S. Wang, Y. Wei, X. Yang, and K.-L. Tsui, “Forecasting tourist arrivals with machine learning and internet search index,” 2017, pp. 4165–4169.
- [14] S. Levantesi and V. Pizzorusso, “Application of Machine Learning to Mortality Modeling and Forecasting,” *Risks*, vol. 7, p. 26, Feb. 2019.
- [15] M. Battifarano, D. DeSmet, A. Madabhushi, and P. Nabar, “Predicting Future Machine Failure from Machine State Using Logistic Regression,” *arXiv:1804.06022 [stat]*, Apr. 2018.
- [16] S. Kullback and R. A. Leibler, “On Information and Sufficiency,” *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, Mar. 1951.
- [17] Z. Tian, “An artificial neural network method for remaining useful life prediction of equipment subject to condition monitoring,” *Journal of Intelligent Manufacturing*, vol. 23, pp. 227–237, 2012.
- [18] P. Malhotra, L. Vig, G. Shroff, and P. Agarwal, “Long Short Term Memory Networks for Anomaly Detection in Time Series,” *Computational Intelligence*, p. 6, 2015.
- [19] R. Sipos, D. Fradkin, F. Moerchen, and Z. Wang, “Log-based predictive maintenance,” in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD ’14*, 2014, pp. 1867–1876.