

Dipartimento di
Ingegneria Gestionale, dell'Informazione e della Produzione



Corso di laurea in
Ingegneria informatica

Classe n. L-8 - Classe delle lauree in Ingegneria
dell'informazione

Interactive AI: quando l'AI riesce a collaborare con l'uomo

Candidata:
Georgina Botello Ortega

Relatore:
Chiar.mo Prof. Angelo Gargantini

Matricola n. 1041818

Anno Accademico
2022/2023

“L'IA interattiva è una tecnologia che consente alle macchine di interagire con gli esseri umani in modo più naturale e intuitivo. Utilizza algoritmi di Natural Language Processing (NLP) e Machine Learning (ML) per comprendere il linguaggio umano e rispondere di conseguenza. Questa tecnologia può essere utilizzata per una varietà di applicazioni, come il servizio clienti, i chatbot, gli assistenti virtuali e altro ancora. Con l'intelligenza artificiale interattiva, è possibile creare interfacce conversazionali in grado di comprendere l'input dell'utente e fornire risposte appropriate.

L'intelligenza artificiale interattiva ha il potenziale per rivoluzionare il modo in cui interagiamo con le macchine, rendendole più reattive e intuitive. Sfruttando algoritmi ML avanzati, può fornire agli utenti un'esperienza coinvolgente adattata alle loro esigenze. Inoltre, l'intelligenza artificiale interattiva può aiutare a ridurre i costi associati al servizio clienti fornendo soluzioni automatizzate per le richieste comuni.”

Scritto da **“Rytr - Best AI Writer, Content Generator & Writing Assistant”**, un assistente virtuale per generare contenuti di qualità. Alla fine della tesi confronterò se tutto ciò che ha scritto corrisponde al vero.

INDICE

1. INTRODUZIONE	7
2. INTRODUZIONE AL MACHINE LEARNING E ALLE IA	
2.1 Cos'è l'IA	9
2.2 Diversi sottocampi dell'IA	10
2.2.1 Cos'è il Machine Learning	11
2.2.2 Deep Learning	12
2.2.3 Neural Networks	12
2.3 Cosa sono l'Interactive AI e l'Interactive Human-centered AI	Error! Bookmark not defined.
2.4 Tipi di algoritmi di Machine Learning	15
2.4.1 I modelli del Machine Learning	16
2.4.2 Reinforcement Machine Learning	17
2.4.3 Algoritmi comuni di apprendimento automatico	17
2.4.4 Vantaggi e svantaggi	18
3. INTERACTIVE AI	
3.1 Error! Bookmark not defined.	
3.2 Black Box Problem	22
3.3 Come creare una Interactive AI	26

3.4 Vantaggi e svantaggi	27
3.5 Applicazioni	29
4. PROCESSO DI APPRENDIMENTO COLLABORATIVO	31
4.1 Interazione Umano - AI per il riconoscimento delle figure geometriche	
5. CONCLUSIONI	37
6. BIBLIOGRAFIA E SITOGRAFIA	40

1. INTRODUZIONE

L'obiettivo di questa tesi è quello di introdurre l'argomento di *Interactive AI*, capire come e perché sia nato questo campo specifico dell'Intelligenza Artificiale, quali benefici apporta e che soluzioni può offrire.

Da un punto di vista antropocentrico, l'IA è quindi vista come uno strumento molto potente e ci si chiede, oltre alle questioni etiche, quali conseguenze potrebbe avere uno strumento che si “controlla da solo”. Questo, infatti, è quello che può e rischia di accadere quando l'uomo non viene consultato nei processi decisionali dell'algoritmo o quando questi risultano incomprensibili.

Sentiamo quindi il bisogno di una IA Interattiva e Centrata sull'uomo (*Human-Centered*), per capire come possiamo collaborare, imparando l'uno dall'altra e aumentando la nostra comprensione globale dei problemi e soluzioni.

Nel quarto capitolo, viene utilizzato un algoritmo in linguaggio Java per il riconoscimento delle figure geometriche, in particolare per distinguere i quadrati dai rettangoli. Il programma richiede all'utente di fornire un dataset contenente le coordinate cartesiane delle figure.

Il modello elabora le coordinate fornite e genera una previsione sul tipo di figura (quadrato o rettangolo) corrispondente. Successivamente, l'utente viene coinvolto nel processo di apprendimento del modello, in quanto gli viene chiesto di confermare o smentire la previsione generata.

Questa interazione con l'utente comporta sia aspetti positivi che negativi. Da un lato, consente di sfruttare l'esperienza umana e le competenze nel riconoscimento delle figure che potrebbero non essere facilmente identificabili ad occhio nudo. D'altro canto, introduce anche la possibilità di errore umano nel processo di apprendimento.

In questo modo, il modello dipende direttamente dalle indicazioni fornite dall'utente, rendendolo una sorta di assistente che si affida alle direttive umane. L'interazione tra l'utente e il modello aiuta ad affinare le capacità di riconoscimento delle figure geometriche, consentendo al modello di apprendere dagli errori e migliorare le sue previsioni.

2. INTRODUZIONE AL MACHINE LEARNING E ALLE IA

2.1 Cos'è l'IA

L'Intelligenza Artificiale (AI) è una tecnologia rivoluzionaria che sta rapidamente trasformando la nostra società, insieme alla nostra economia e lavoro. Alcuni esempi popolari di applicazioni AI sono i social network, le auto senza conducente, le chatbots, gli assistenti vocali, i motori di ricerca su internet, i trader di borsa robotici, ecc. Questi sistemi possono essere incorporati in macchine fisiche o essere software autonomi che agiscono nello spazio digitale. La loro sorprendente capacità di cambiare il nostro stile di vita richiederà alla società e ai politici di comprendere e regolamentare le funzioni e i limiti di questi dispositivi [15].

Se vogliamo dare una definizione più precisa, possiamo dire che l'intelligenza artificiale (AI) è la capacità delle macchine di replicare o potenziare l'intelletto umano, come il ragionamento e l'apprendimento dall'esperienza. L'intelligenza artificiale è stata utilizzata in programmi informatici per anni, ma ora viene applicata a molti altri prodotti e servizi. Ad esempio, alcune fotocamere digitali possono determinare quali oggetti sono presenti in un'immagine utilizzando software di intelligenza artificiale. Inoltre, gli esperti prevedono molti altri usi innovativi per l'intelligenza artificiale in futuro.

L'IA utilizza tecniche derivate dalla teoria delle probabilità, dall'economia e dalla progettazione degli algoritmi per risolvere problemi pratici. Inoltre, il campo dell'IA attinge dalla scienza informatica, dalla matematica, dalla psicologia e dalla linguistica. La scienza informatica fornisce strumenti per progettare e costruire algoritmi, mentre la matematica offre strumenti per modellare e risolvere i problemi di ottimizzazione risultanti.

Sebbene il concetto di intelligenza artificiale esista fin dal XIX secolo, quando Alan Turing ha proposto per la prima volta un "gioco dell'imitazione" per valutare l'intelligenza delle macchine, è diventato fattibile da realizzare solo negli ultimi decenni grazie alla maggiore disponibilità di potenza di calcolo e dati per addestrare i sistemi di intelligenza artificiale.

Per capire l'idea alla base dell'AI, è necessario riflettere su ciò che distingue l'intelligenza umana da quella delle altre creature: la nostra capacità di imparare dall'esperienza e applicare queste lezioni a nuove situazioni. Possiamo farlo grazie alla nostra avanzata potenza cerebrale; abbiamo più neuroni di qualsiasi altra specie animale.

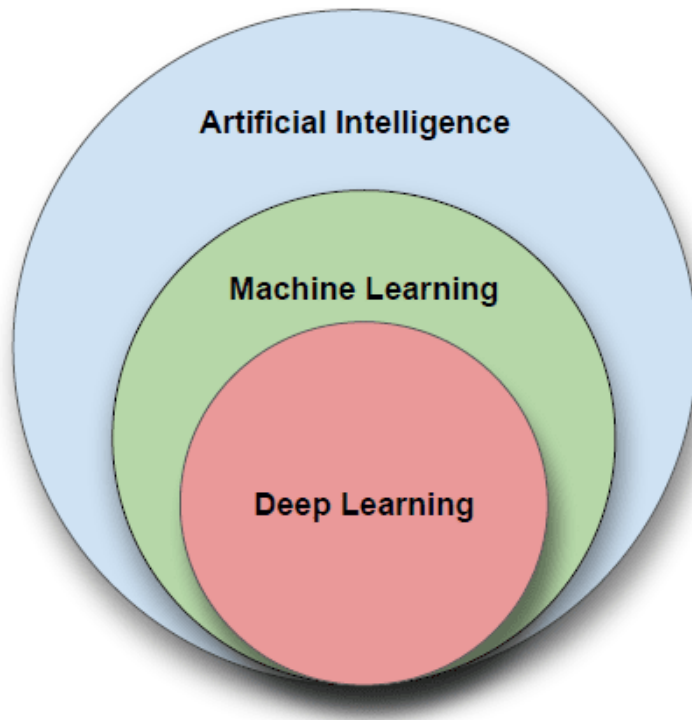
I computer attuali non raggiungono la rete neurale biologica umana, nemmeno lontanamente. Ma hanno un vantaggio significativo rispetto a noi: la loro capacità di analizzare grandi quantità di dati ed esperienze molto più velocemente di quanto gli esseri umani potrebbero mai sperare.

L'IA consente di concentrarsi sui compiti più critici e prendere decisioni migliori basate sui dati acquisiti relativi a un caso d'uso. Può essere utilizzata per compiti complessi, come prevedere le necessità di manutenzione, individuare frodi con carte di credito e trovare il percorso migliore per un camion di consegna. In altre parole, l'IA può automatizzare molti processi aziendali, consentendoci di concentrarci sul core business.

2.2 Diversi sottocampi dell'IA

L'intelligenza artificiale è il campo più in voga nell'informatica. Tuttavia, con tutte le nuove tecnologie e ricerche, sta crescendo così rapidamente che può essere ambiguo capire cosa sia. Inoltre, ci sono molti settori diversi all'interno dell'AI, ognuno con i suoi specifici algoritmi. Pertanto, è essenziale sapere che l'IA non è un singolo campo ma una combinazione di vari campi. L'intelligenza artificiale (IA) è il termine generale per la capacità di far svolgere ai computer

compiti che richiedono intelligenza se svolti da esseri umani. L'IA può essere suddivisa in due grandi campi, Machine Learning (ML) e Reti Neurali (NN). Entrambi sono sottocampi dell'intelligenza artificiale e ognuno ha i suoi metodi e algoritmi per aiutare a risolvere i problemi.



2.2.1 Cos'è il Machine Learning

Il Machine Learning (ML) permette ai computer di imparare dai dati e dall'esperienza al fine di migliorare le loro prestazioni in determinati compiti o processi decisionali. Il ML utilizza la statistica e la teoria della probabilità a questo scopo. Il Machine Learning utilizza algoritmi per analizzare i dati, imparare da essi e prendere decisioni senza programmazione esplicita. Gli algoritmi di Machine Learning vengono spesso categorizzati come supervisionati o non supervisionati. Gli algoritmi supervisionati possono applicare ciò che è stato appreso in passato a nuovi insiemi di dati, mentre gli

algoritmi non supervisionati possono dedurre informazioni dagli insiemi di dati. Gli algoritmi di Machine Learning sono progettati per cercare di stabilire relazioni lineari e non lineari in un determinato insieme di dati. Questa impresa viene raggiunta attraverso metodi statistici utilizzati per addestrare l'algoritmo a classificare o predire da un insieme di dati.

2.2.2 Deep Learning

Il deep learning è una sottoarea dell'apprendimento automatico che utilizza reti neurali artificiali multi-strato per fornire un'accuratezza all'avanguardia nella rilevazione degli oggetti, nel riconoscimento del linguaggio parlato e nella traduzione linguistica. Il deep learning è una tecnologia fondamentale dietro le auto senza conducente e consente l'analisi automatica di grandi quantità di dati complessi, ad esempio il riconoscimento dei volti delle persone che compaiono in un'immagine o in un video.

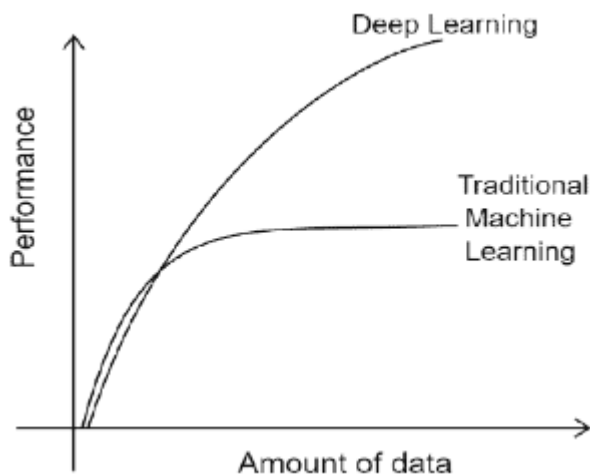
2.2.3 Neural Networks

Le reti neurali sono ispirate ai neuroni biologici presenti nel cervello umano e sono composte da strati di nodi connessi chiamati "neuroni", che contengono funzioni matematiche per elaborare i dati in ingresso e prevedere un valore di output. Le reti neurali artificiali imparano attraverso l'esempio, in modo simile a come gli esseri umani imparano dai genitori, dagli insegnanti e dai pari. Esse sono composte almeno da tre strati: uno strato di input, strati nascosti e uno strato di output. Ogni strato contiene nodi (anche chiamati neuroni) che hanno input pesati per calcolare l'output.

Le prestazioni dei modelli tradizionali di apprendimento automatico raggiungono un limite e l'aggiunta di ulteriori dati non aiuta a migliorarne le prestazioni. I modelli di deep learning continuano a migliorare le prestazioni con l'aumento dei dati.

Questi campi hanno algoritmi diversi, a seconda del caso d'uso. Ad esempio, per l'apprendimento automatico abbiamo alberi decisionali, foreste casuali, boosting, macchine a vettori di supporto (SVM), vicini più vicini (kNN) e altri. Per le reti neurali, abbiamo reti neurali convoluzionali (CNN), reti neurali ricorrenti (RNN), reti neurali LSTM (long short-term memory) e altro ancora.

Tuttavia, classificare l'IA in base alla sua potenza e alle sue capacità significherebbe suddividerla ulteriormente in "IA stretta" e "IA generale". L'IA stretta riguarda il far svolgere alle macchine una singola attività in modo eccellente, come il riconoscimento delle immagini o il gioco degli scacchi. L'IA generale indica dispositivi in grado di fare tutto ciò che gli esseri umani possono fare e di più. Le ricerche odierne si concentrano sull'IA stretta, ma molti ricercatori vorrebbero che l'apprendimento automatico raggiungesse alla fine l'IA generale.



2.3 Cosa sono l'Interactive AI e l'Interactive Human-centered AI?

Per poter rispondere a questa domanda, è necessario fare una premessa. Abbiamo visto che le potenzialità di questo strumento sono ancora indefinite, stiamo ancora scoprendo ogni giorno e non possiamo delimitare le possibilità dell'IA. Per questo ci sono molti dubbi e anche molte paure. Dopotutto, essendo uno strumento molto potente necessita di essere usato un certo grado di etica per poter sfruttare appieno i suoi benefici.

Proprio per questo è nata un'idea di IA centrata sull'essere umano, in cui le macchine collaborano con le persone per migliorare il controllo e l'agenzia umana. Questo approccio, chiamato Intelligenza Artificiale Centrata sull'Umano (HCAI), mira ad amplificare l'agenzia umana invece di sostituirla. Si sottolinea l'importanza di immaginare, parlare e progettare i sistemi di IA in modo che le persone mantengano il controllo della tecnologia e traggano beneficio dalle informazioni che consentono loro di porre domande migliori e prendere decisioni più audaci.

Dall'altro lato, un'Intelligenza Artificiale Interattiva Centrata sull'Uomo (Interactive Human Centered Artificial Intelligence) è un tipo di Intelligenza Artificiale che consente l'esplorazione e la manipolazione interattiva in tempo reale ed è progettata con uno scopo chiaro a beneficio dell'essere umano, mantenendo trasparenza su chi ha il controllo dei dati e degli algoritmi.

In generale si hanno queste quattro linee guida:

- Gli individui possono interagire in tempo reale con gli algoritmi, i modelli e i dati e possono manipolare e controllare tutti i parametri rilevanti.
- Si spiega quali rischi l'IA comporta per gli individui e per la società a livello generale.

- È visibile chi ha il controllo dell'intelligenza artificiale, chi ha il potere sui dati, i modelli e gli algoritmi.
- Per verificare che un'IA sia centrata sull'essere umano, è essenziale conoscere gli obiettivi di progettazione e chi ha il controllo dei dati e degli algoritmi.

2.4 Tipi di algoritmi di Machine Learning

Il machine learning, come abbiamo detto nell'introduzione, è una disciplina che si occupa dell'utilizzo di dati e algoritmi per imitare il processo di apprendimento umano e migliorare progressivamente la precisione delle previsioni.

Negli ultimi anni, grazie ai progressi tecnologici in termini di capacità di archiviazione e potenza di calcolo, sono state realizzate applicazioni innovative basate sul machine learning, come il sistema di raccomandazione di Netflix e le auto a guida autonoma. Il machine learning svolge un ruolo importante nel campo della data science, in cui gli algoritmi vengono addestrati a effettuare classificazioni, previsioni e scoprire informazioni rilevanti nei progetti di data mining. Queste informazioni supportano la presa di decisioni aziendali e influenzano le metriche di crescita. Con l'aumento dei dati disponibili, la richiesta di data scientist, esperti nella gestione e analisi dei dati, è in costante crescita.

Il machine learning è un processo composto da tre componenti principali:

1. Processo decisionale: Gli algoritmi di machine learning vengono utilizzati per fare previsioni o classificazioni. Basandosi su dati di input, l'algoritmo produce una stima di un pattern presente nei dati.

2. Funzione di errore: Una funzione di errore valuta la precisione delle previsioni del modello. Se sono disponibili esempi noti, la funzione di errore confronta i risultati per valutare l'accuratezza del modello.

3. Processo di ottimizzazione del modello: Se il modello può essere migliorato per adattarsi meglio ai punti dati nel set di addestramento, i pesi vengono regolati per ridurre la discrepanza tra gli esempi noti e le stime del modello. L'algoritmo ripete questo processo di "valutazione e ottimizzazione", aggiornando autonomamente i pesi finché non viene raggiunta una soglia di accuratezza desiderata.

In sintesi, il machine learning utilizza un processo decisionale per fare previsioni, una funzione di errore per valutare l'accuratezza delle previsioni e un processo di ottimizzazione del modello per migliorare la precisione del modello stesso.

2.4.1 I modelli del Machine learning

I modelli di machine learning si suddividono in tre categorie principali:

1. Supervised machine learning (apprendimento supervisionato): Utilizza dataset etichettati per addestrare algoritmi a classificare dati o prevedere risultati. Adatto per problemi in cui si dispone di dati etichettati e si desidera creare modelli di previsione.

2. Unsupervised machine learning (apprendimento non supervisionato): Analizza e raggruppa dataset non etichettati per scoprire modelli o raggruppamenti nascosti. Utile per l'analisi esplorativa dei dati e la scoperta di informazioni senza l'utilizzo di etichette predefinite.

3. Semi-supervised learning (apprendimento semi-supervisionato): Combina elementi dell'apprendimento supervisionato e non supervisionato. Utilizza un piccolo set di dati etichettati insieme a un set più grande di dati non etichettati per guidare la classificazione e l'estrazione delle caratteristiche.

Queste categorie rappresentano approcci diversi nell'addestramento degli algoritmi di machine learning e vengono utilizzate in base alle caratteristiche dei dati disponibili e agli obiettivi dell'applicazione.

2.4.2 Reinforcement Machine Learning

Il reinforcement machine learning è un modello di apprendimento automatico simile all'apprendimento supervisionato, ma l'algoritmo non viene addestrato utilizzando dati di esempio. Questo modello impara attraverso il metodo del tentativo ed errore. Una sequenza di risultati positivi viene rinforzata per sviluppare la migliore raccomandazione o politica per un dato problema. Un buon esempio di questo approccio è rappresentato dal sistema IBM Watson che ha vinto la sfida Jeopardy! nel 2011. Il sistema utilizzava l'apprendimento per rinforzo per imparare quando rispondere (o formulare una domanda, nel caso del gioco), quale casella selezionare sulla scacchiera e quanto puntare, specialmente sui "double" giornalieri.

2.4.3 Algoritmi comuni di apprendimento automatico

Esistono diversi algoritmi di apprendimento automatico comunemente utilizzati, questi includono:

Reti neurali: Le reti neurali simulano il funzionamento del cervello umano, con un gran numero di nodi di elaborazione collegati. Le reti neurali sono in grado di riconoscere modelli e svolgono un ruolo importante in applicazioni come la traduzione automatica, il riconoscimento delle immagini, il riconoscimento del parlato e la generazione di immagini.

Regressione lineare: Questo algoritmo viene utilizzato per prevedere valori numerici, basandosi su una relazione lineare tra diversi valori. Ad esempio, questa tecnica potrebbe essere utilizzata per prevedere i prezzi delle case basandosi sui dati storici dell'area.

Regressione logistica: Questo algoritmo di apprendimento supervisionato effettua previsioni per variabili di risposta categoriche, come risposte "sì/no" a delle domande. Può essere utilizzato per applicazioni come la classificazione dello spam e il controllo di qualità in una linea di produzione.

Clustering: Utilizzando l'apprendimento non supervisionato, gli algoritmi di clustering possono identificare modelli nei dati in modo da poterli raggruppare. I computer possono aiutare gli scienziati dei dati identificando differenze tra gli elementi di dati che gli esseri umani potrebbero aver trascurato.

Alberi decisionali: Gli alberi decisionali possono essere utilizzati sia per prevedere valori numerici (regressione) che per classificare i dati in categorie. Gli alberi decisionali utilizzano una sequenza ramificata di decisioni collegate che possono essere rappresentate con un diagramma ad albero. Uno dei vantaggi degli alberi decisionali è che sono facili da convalidare e controllare, a differenza della "scatola nera" delle reti neurali.

Foreste casuali: In una foresta casuale, l'algoritmo di apprendimento automatico prevede un valore o una categoria combinando i risultati di un certo numero di alberi decisionali.

2.4.4 Vantaggi e svantaggi:

I vantaggi derivanti dall'uso e sviluppo di questi algoritmi possono essere già constatati oggi da tutti noi. Ci sono innumerevoli casi d'uso, tra cui :

Riconoscimento del parlato: È anche noto come riconoscimento automatico del parlato (ASR), riconoscimento del parlato da parte del computer o conversione del parlato in testo ed è una capacità che utilizza l'elaborazione del linguaggio naturale (NLP) per tradurre il parlato umano in un formato scritto. Molti dispositivi mobili incorporano il riconoscimento del parlato nei loro sistemi per condurre ricerche vocali, ad esempio Siri, o migliorare l'accessibilità per la scrittura di messaggi.

Servizio clienti: I chatbot online stanno sostituendo gli agenti umani lungo il percorso del cliente, cambiando il modo in cui concepiamo l'interazione con i clienti su siti web e piattaforme di social media. I chatbot rispondono alle domande frequenti (FAQ) su argomenti come la spedizione o forniscono consigli personalizzati, promuovendo la vendita incrociata di prodotti o suggerendo le taglie agli utenti. Gli esempi includono agenti virtuali su siti di e-commerce, bot di messaggistica utilizzando Slack e Facebook Messenger e compiti di solito eseguiti da assistenti virtuali e assistenti vocali.

Visione artificiale: Questa tecnologia di intelligenza artificiale consente ai computer di ottenere informazioni significative da immagini digitali, video e altre informazioni visive e di intraprendere l'azione appropriata. Potenziata da reti neurali convoluzionali, la visione artificiale ha applicazioni nel riconoscimento delle foto sui social media, nell'imaging radiologico in ambito sanitario e nelle auto a guida autonoma nel settore automobilistico.

Motori di raccomandazione: Utilizzando i dati sul comportamento di consumo passato, gli algoritmi di intelligenza artificiale possono aiutare a scoprire tendenze dei dati che possono essere utilizzate per sviluppare strategie di vendita incrociata più efficaci. Questo approccio viene utilizzato dai rivenditori online per fornire raccomandazioni di prodotti pertinenti ai clienti durante il processo di checkout.

Trading azionario automatizzato: Progettate per ottimizzare i portafogli azionari, le piattaforme di trading ad alta frequenza basate sull'intelligenza artificiale effettuano migliaia o addirittura milioni di negoziazioni al giorno senza intervento umano.

Individuazione delle frodi: Le banche e altre istituzioni finanziarie possono utilizzare l'apprendimento automatico per individuare transazioni sospette. L'apprendimento supervisionato può addestrare un modello utilizzando informazioni su transazioni fraudolente note. La rilevazione delle anomalie può

identificare transazioni che sembrano atipiche e che richiedono ulteriori indagini.

Come per tutte le invenzioni, e tutti gli strumenti, ci sono anche possibili “effetti collaterali” che derivano dal modo e soprattutto dalle finalità che ha chi usa quello strumento o dall’inevitabile cambiamento che questo comporta (pur essendo estremamente positivo, il cambiamento spaventerà sempre). Sono quindi molte le preoccupazioni etiche riguardo alle tecnologie di intelligenza artificiale. Alcuni di questi punti includono:

1. Singolarità tecnologica: La possibilità che l'intelligenza artificiale superi l'intelligenza umana nel futuro prossimo solleva domande interessanti. Ad esempio, nei sistemi autonomi come le auto senza conducente, chi è responsabile in caso di incidenti? Dovremmo sviluppare veicoli autonomi o limitare questa tecnologia a veicoli semi-autonomi? Questi sono dibattiti etici che si stanno svolgendo mentre nuove tecnologie di intelligenza artificiale si sviluppano.

2. Impatto dell'IA sui posti di lavoro: La percezione pubblica dell'intelligenza artificiale spesso si concentra sulla perdita di posti di lavoro. Tuttavia, con ogni nuova tecnologia innovativa, la domanda di ruoli lavorativi specifici si sposta. L'intelligenza artificiale cambierà la domanda di lavoro verso altre aree, ad esempio nel gestire i sistemi di intelligenza artificiale stessi o nell'affrontare problemi complessi all'interno di settori interessati dai cambiamenti nella domanda di lavoro.

3. Privacy: La privacy viene spesso discussa in termini di protezione e sicurezza dei dati. La legislazione come il GDPR in Europa e il CCPA negli Stati Uniti sono stati introdotti per garantire maggior controllo ai cittadini sui propri dati personali. Le preoccupazioni sulla privacy hanno spinto le aziende a investire nella sicurezza dei dati per eliminare vulnerabilità e rischi di sorveglianza e attacchi informatici.

4. Bias e discriminazione: La presenza di bias e discriminazione in diversi sistemi di machine learning solleva domande etiche sull'uso dell'intelligenza artificiale. Come possiamo proteggerci dal bias e dalla discriminazione quando i dati di addestramento stessi possono essere generati da processi umani prevenuti da bias? È necessario affrontare il bias e la discriminazione in applicazioni come il riconoscimento facciale e gli algoritmi dei social media.

5. Responsabilità: Attualmente non esistono leggi significative per regolamentare le pratiche dell'IA, quindi non ci sono meccanismi di applicazione per garantire la pratica etica dell'IA. Per colmare questa lacuna, sono emerse linee guida etiche come collaborazione tra eticisti e ricercatori per governare la costruzione e la distribuzione dei modelli di intelligenza artificiale nella società.

3. INTERACTIVE AI

3.1 Definizione

Abbiamo già introdotto nel capitolo precedente cosa sia una IA Interattiva, ma una definizione più tecnica di cosa è sarebbe:

Un'Intelligenza Artificiale Interattiva che consente l'esplorazione e la manipolazione interattiva in tempo reale ed è progettata con uno scopo chiaro per il beneficio umano, mantenendo la trasparenza su chi ha il controllo dei dati e degli algoritmi.

In breve le proprietà che questa deve avere sono dunque:

- Le persone possono interagire in tempo reale con gli algoritmi, i modelli e i dati e possono manipolare e controllare tutti i parametri rilevanti.
- Spiega quali rischi l'IA rappresenta per gli individui e per la società.
- È visibile chi ha il controllo dell'intelligenza artificiale, chi detiene il potere sui dati, i modelli e gli algoritmi.

Per verificare che un'IA sia centrata sull'essere umano, è essenziale conoscere gli obiettivi di progettazione e chi ha il controllo dei dati e degli algoritmi.

Per capire perché sia nato questo filone di IA, è necessario capire i problemi che sono sorti quando abbiamo iniziato a sviluppare l'IA.

3.2 Black Box Problem

Il problema noto come “The Black Box Problem” è ormai noto e riassunto in poche parole: “questo problema scaturisce quando gli algoritmi non forniscono alcuna giustificazione o motivazione per i loro risultati. Gli utenti hanno quindi bisogno di comprenderli per poter prendere le giuste decisioni.”

Sebbene questo sia un problema abbastanza rilevante, soprattutto per chi deve poi decidere in base agli output di questi algoritmi, in realtà costituisce solo parte o per meglio dire metà, dell'intero problema per l'analisi dei dati. C'è

un'altra "black box" nell'equazione: quello della cognizione umana. Dal punto di vista dell'algoritmo, la mente umana è un'altra "black box" altrettanto (o forse anche più) difficile da interpretare.

Come può quindi l'essere umano "spiegarsi" all'algoritmo, in modo che quest'ultimo possa rispondere ai suoi processi di pensiero interni, così come ai suoi obiettivi, motivazioni, conoscenze esperte del dominio e intenzioni? Come può la macchina imparare dall'interazione con l'utente?

Ci ritroviamo così ad avere non solo una, ma ben due variabili nel nostro problema:

XAI □ IA SPIEGABILE (AI explainable)

XCI □ INTELLIGENZA COGNITIVA SPIEGABILE (IC explainable)

Per poter risolvere queste incognite e creare un tale apprendimento condiviso ibrido tra umani e AI, ci sono cinque sfide significative che devono essere affrontate:

Sfida 1: Cos'è il Loop (ciclo)?

"L'essere umano nel Loop" descrive sistemi interattivi che sono comunemente progettati in modo che gli algoritmi analitici consultino occasionalmente gli esseri umani per un feedback esperto.

"La macchina nel Loop", al contrario è la situazione in cui l'essere umano è il principale responsabile del processo di analisi, consultando occasionalmente la macchina per suggerimenti o assistenza in problemi che richiedono abilità computazionali.

L'IA non dovrebbe solo assistere l'utente, ma dovrebbe anche migliorare la sua comprensione degli obiettivi e dei processi dell'utente al fine di fornire una migliore assistenza nel tempo.

L'obiettivo è formare un ciclo di apprendimento cooperativo. Perché questo ciclo di apprendimento cooperativo funzioni in modo ottimale, sia l'essere umano che l'IA devono comprendere reciprocamente l'altro.

A questa si aggiungono altre quattro ulteriori sfide di comunicazione:

1. Estrarre la cognizione dall'essere umano.
2. Presentare quella cognizione in modo significativo all'IA.
3. Estrarre informazioni dall'IA.
4. Presentare quelle informazioni in modo significativo all'essere umano.

Sfida 2: Esporre la Cognizione

Estrarre lo stato cognitivo dell'essere umano è una sfida significativa che deve essere affrontata.

Sfida 3: Inserimento Cognitivo nell'IA

Convertire la cognizione esternalizzata dell'essere umano in una forma comprensibile per una macchina. Le interazioni ad alto livello eseguite dall'essere umano devono essere trasformate in un insieme di parametri e valori che riflettono le intenzioni alla base delle interazioni.

Sfida 4: Output contestuale dell'AI

La sfida 4 riguarda l'output contestuale dell'IA. L'obiettivo dell'Intelligenza Artificiale Esplicabile (XAI) è quello di fornire informazioni interpretabili dagli modelli di intelligenza artificiale agli esseri umani, in modo che possano capire e fidarsi dei risultati del modello. Questo può essere integrato nei sistemi di analisi interattiva centrati sull'essere umano per migliorare la comprensione dei dati da parte dell'utente e facilitare il processo di analisi. (Nel sistema StarSPIRE le informazioni estratte dall'IA vengono contestualizzate e integrate nello spazio di lavoro visuale dell'analista in modi diversi, come la visualizzazione dei documenti in base alla similarità e l'evidenziazione dei termini rilevanti. Questo output contestuale aiuta l'analista a individuare

informazioni utili e a ottenere un'idea di ciò che l'IA ha identificato come importante.)

L'obiettivo è fornire informazioni rilevanti e interpretabili in modo da supportare la comprensione umana e costruire fiducia nel modello di intelligenza artificiale.

Sfida 5: Comprendere la Macchina

In breve, la sfida consiste nel combinare l'apprendimento umano incrementale con l'approccio batch dell'apprendimento automatico al fine di incorporare l'intelligenza artificiale (IA) nel processo cognitivo umano. È importante progettare algoritmi che apprendano gradualmente attraverso il feedback umano e forniscono risultati incrementali, in modo che l'IA possa essere accettata dagli esseri umani. La collaborazione tra esseri umani e IA richiede precauzioni per evitare conclusioni affrettate, creare collegamenti supportati da prove e non dedurre prematuramente gli interessi umani. (Il sistema StarSPIRE adotta un approccio incrementale, aggiornando continuamente i modelli, il layout e i documenti visualizzati in modo graduale. Gli utenti aprono i documenti uno alla volta e i loro interessi vengono aggiornati lentamente senza cambiamenti drastici. Questo approccio mira a mantenere il modello mentale dell'utente umano e evitare confusione.) La stretta interazione tra l'IA che apprende dalle interazioni umane e gli esseri umani che apprendono dai dati dimostra una relazione di co-apprendimento, in cui entrambi contribuiscono con le proprie abilità per risolvere il problema in modo collaborativo.

3.3 Come creare una interactive AI

La creazione di un' IA interattiva richiede una serie di passaggi e considerazioni. Ecco una panoramica generale del processo:

1. Definizione degli obiettivi: Prima di iniziare a sviluppare un'IA interattiva, è fondamentale stabilire chiari obiettivi e finalità. Ciò include identificare le aree in cui si desidera che l'IA interagisca con gli utenti, definire le funzionalità e le capacità desiderate e comprendere come l'IA può migliorare l'esperienza utente o risolvere specifici problemi.
2. Raccolta e preparazione dei dati: L'IA interattiva richiede dati di input per apprendere e fornire risposte pertinenti. È necessario raccogliere dati pertinenti e prepararli per l'addestramento dell'IA. Ciò può includere la pulizia dei dati, l'etichettatura, la normalizzazione e la creazione di set di dati di addestramento, validazione e test.
3. Progettazione dell'architettura dell'IA: L'architettura dell'IA interattiva dipende dalla specifica applicazione e dalle esigenze dell'utente. È possibile utilizzare diversi approcci, come reti neurali, algoritmi di apprendimento automatico o metodi basati su regole. La scelta dipende dalle caratteristiche del problema da risolvere e dalle disponibilità di dati.
4. Sviluppo dei modelli e dell'algoritmo: Una volta progettata l'architettura, è necessario sviluppare i modelli di intelligenza artificiale e gli algoritmi che consentiranno all'IA di interagire con gli utenti in tempo reale. Ciò può richiedere l'addestramento di modelli di apprendimento automatico utilizzando i dati raccolti, l'ottimizzazione degli algoritmi per massimizzare l'interazione utente-IA e la valutazione delle prestazioni del modello.
5. Implementazione dell'interfaccia utente: L'IA interattiva richiede un'interfaccia utente che consenta agli utenti di comunicare con l'IA e ricevere risposte o risultati. L'interfaccia può essere realizzata attraverso applicazioni web, chatbot, app mobili o altri mezzi interattivi a seconda del contesto e delle preferenze degli utenti.
6. Valutazione e iterazione: Dopo l'implementazione dell'IA interattiva, è importante valutare le prestazioni e l'efficacia dell'IA rispetto agli obiettivi

stabiliti inizialmente. La raccolta dei feedback degli utenti, l'analisi delle metriche di successo e il monitoraggio continuo delle prestazioni consentono di identificare eventuali problemi o aree di miglioramento. In base ai risultati, possono essere apportate modifiche e iterazioni per ottimizzare ulteriormente l'IA.

È importante sottolineare che la creazione di un' IA interattiva è un processo complesso che richiede competenze multidisciplinari, tra cui scienza dei dati, ingegneria del software, progettazione dell'interfaccia utente.

3.4 Vantaggi e svantaggi

L'Intelligenza Artificiale Interattiva presenta una serie di vantaggi e svantaggi. Ecco un elenco dei principali vantaggi:

1. **Migliore supporto decisionale:** L'IA interattiva può fornire assistenza in tempo reale e suggerimenti basati sui dati per supportare le decisioni umane, migliorando la precisione e l'efficacia delle scelte.
2. **Esperienza personalizzata:** Grazie all'interazione con l'utente, l'IA interattiva può adattarsi alle preferenze, ai bisogni e al contesto specifico dell'individuo, offrendo un'esperienza personalizzata e rilevante.
3. **Ottimizzazione dei processi:** L'IA interattiva può analizzare i dati in modo rapido ed efficiente, consentendo di ottimizzare i processi, migliorare l'efficienza operativa e ridurre i costi.
4. **Apprendimento continuo:** Attraverso l'interazione con gli utenti, l'IA interattiva può apprendere e adattarsi nel tempo, migliorando la sua comprensione dei bisogni e fornendo una migliore assistenza nel tempo.
5. **Miglior comprensione dei dati:** L'IA interattiva può aiutare gli utenti a esplorare e comprendere meglio grandi quantità di dati complessi, identificando modelli, tendenze o relazioni che potrebbero sfuggire all'occhio umano.

Svantaggi:

1. Dipendenza dai dati di addestramento: L'IA interattiva richiede dati di addestramento di alta qualità per ottenere risultati accurati. La mancanza o la presenza di dati non rappresentativi può portare a risultati errati o parziali.
2. Rischi di privacy e sicurezza: L'interazione con l'IA interattiva può implicare la condivisione di dati personali e sensibili. Ciò può comportare rischi di violazione della privacy e della sicurezza delle informazioni.
3. Dipendenza tecnologica: L'IA interattiva può portare a una maggiore dipendenza dalla tecnologia, con la possibilità di perdere alcune abilità o competenze umane nel processo.
4. Complessità e opacità: Gli algoritmi di intelligenza artificiale possono essere complessi e opachi, rendendo difficile per gli utenti comprendere come vengono presi determinati risultati o decisioni.
5. Possibilità di bias: L'IA interattiva può essere soggetta a bias o discriminazione se i dati di addestramento contengono pregiudizi o disparità presenti nella società umana.

È importante valutare attentamente i vantaggi e gli svantaggi dell'IA interattiva in base al contesto specifico di utilizzo e garantire un uso etico e responsabile delle tecnologie AI per massimizzare i benefici e mitigare i rischi potenziali.

3.5 Applicazioni

L'Intelligenza Artificiale Interattiva trova applicazione in diversi settori e contesti. Alcune delle sue applicazioni includono:

1. Assistenza sanitaria: L'IA interattiva può essere utilizzata per migliorare la diagnosi medica, fornire suggerimenti di trattamento personalizzati e supportare

la cura dei pazienti. Ad esempio, può aiutare i medici a interpretare immagini diagnostiche come raggi X o risonanze magnetiche e a fornire raccomandazioni terapeutiche basate sui dati dei pazienti.

2. Servizi clienti e supporto: I chatbot interattivi basati sull'IA vengono utilizzati per fornire assistenza automatizzata ai clienti, rispondendo alle domande comuni, risolvendo problemi e fornendo informazioni personalizzate.

3. Automazione industriale: L'IA interattiva può essere applicata per monitorare e ottimizzare processi industriali complessi. Ad esempio, può essere utilizzata per rilevare anomalie o guasti nelle macchine, migliorare l'efficienza energetica e ottimizzare la produzione.

4. Veicoli autonomi: L'IA interattiva è un componente chiave dei veicoli autonomi, consentendo loro di percepire l'ambiente circostante, prendere decisioni sulla guida e interagire con i passeggeri. Ciò include funzionalità come il rilevamento degli ostacoli, il riconoscimento dei segnali stradali e la navigazione autonoma.

5. Apprendimento automatico: L'IA interattiva può essere utilizzata per migliorare i processi di apprendimento automatico, consentendo agli utenti di esplorare e manipolare i dati in tempo reale, fornendo feedback sulle predizioni dell'algoritmo e collaborando nell'affinamento dei modelli.

6. Finanza: L'IA interattiva viene utilizzata nel settore finanziario per l'analisi dei dati, la previsione dei mercati, il rilevamento delle frodi e la gestione del rischio. Può supportare gli investitori nelle decisioni di investimento e aiutare le istituzioni finanziarie a migliorare la gestione dei portafogli.

Queste sono solo alcune delle molte possibili applicazioni dell'IA interattiva. La sua versatilità consente di adattarsi a diversi settori e scopi, offrendo potenzialmente vantaggi significativi in termini di efficienza, precisione e personalizzazione delle esperienze umane.

4. PROCESSO DI APPRENDIMENTO COLLABORATIVO:

4.1 Interazione Umano-AI per il riconoscimento delle figure geometriche

Come abbiamo accennato nell'introduzione, presentiamo un' applicazione che mostra una serie di figure geometriche (rettangoli o quadrati) e chiede all'utente di identificarle. Il programma utilizza Weka per fare delle previsioni sulle figure basandosi sui dati forniti dall'utente e li salva in un file ARFF. Weka è un popolare framework open source per l'apprendimento automatico e l'analisi dei dati, fornisce una vasta gamma di algoritmi di apprendimento automatico, tecniche di data mining e strumenti per la visualizzazione dei dati. È ampiamente utilizzato per analizzare dati, costruire modelli predittivi e risolvere problemi di classificazione, regressione, clustering e associazione.

Il nome "Weka" deriva da "Waikato Environment for Knowledge Analysis", poiché è stato sviluppato presso l'Università di Waikato, in Nuova Zelanda.

I passi principali del programma sono:

1. Creazione dell'interfaccia grafica:

Il codice utilizza la libreria Swing per creare una finestra (JFrame) chiamata "Figure" che verrà utilizzata per mostrare le figure geometriche. La finestra viene impostata come visibile e le sue dimensioni vengono impostate a 500x400 pixel.

2. Generazione delle figure:

Il codice utilizza un ciclo while per generare una serie di figure geometriche. Ad ogni iterazione del ciclo, viene generata una figura casuale (rettangolo o

quadrato) utilizzando la classe Rectangle2D e viene disegnata nella finestra utilizzando il metodo paintComponent().

3. Acquisizione della risposta dell'utente:

Dopo aver disegnato la figura, il programma chiede all'utente di identificarla. L'utente può inserire "1" se la figura è un rettangolo o "2" se è un quadrato. La risposta viene acquisita utilizzando l'oggetto Scanner.

4. Salvataggio dei dati:

Vengono salvati i dati relativi alla figura e alla risposta dell'utente in un file ARFF chiamato "Dataset.arff". Viene utilizzato un oggetto BufferedWriter per scrivere i dati nel file.

5. Previsione utilizzando Weka:

Viene creato un oggetto Weka chiamato testing e viene chiamato il metodo Call() per effettuare la previsione sulla figura corrente utilizzando il classificatore Naive Bayes addestrato.

6. Pulizia dell'applet e iterazione:

Dopo aver salvato i dati e effettuato la previsione, l'applet viene pulito rimuovendo tutti i componenti dalla finestra e viene richiamato il metodo repaint() per preparare la finestra alla successiva iterazione del ciclo.

7. Ripetizione del ciclo:

Il ciclo continua fino a quando non vengono generate 20 figure e l'utente non ha fornito la sua risposta per ognuna di esse.

In breve questo programma carica un dataset di addestramento, addestra un classificatore Naive Bayes con il dataset, carica un dataset di test, classifica un'istanza di test utilizzando il classificatore addestrato e restituisce la previsione fatta dal classificatore. L'obiettivo principale è utilizzare

l'apprendimento automatico per classificare nuove istanze di dati in base ai modelli appresi dal classificatore durante l'addestramento.

Primo step: definiamo gli attributi delle coordinate e della Ground truth in un file di tipo arff. Questo verrà usato per registrare le risposte date dall'utente associate alle feature del dataset e di conseguenza anche per addestrare il modello.

```
@RELATION Figure

@ATTRIBUTE vertexAx REAL
@ATTRIBUTE vertexAy REAL
@ATTRIBUTE vertexBx REAL
@ATTRIBUTE vertexBy REAL
@ATTRIBUTE vertexCx REAL
@ATTRIBUTE vertexCy REAL
@ATTRIBUTE vertexDx REAL
@ATTRIBUTE vertexDy REAL
@ATTRIBUTE class {Square, Rectangle}

@DATA
```

Secondo step: lettura del dataset. Per leggere il dataset usiamo la funzione Arffloader. Creiamo un oggetto di tipo Arff a cui assegniamo il dataset letto

```
ArffLoader loader = new ArffLoader();
loader.setFile(new File("Dataset.arff"));
Instances structure = loader.getStructure();
structure.setClassIndex(structure.numAttributes() - 1);
```

Terzo step: costruzione del classificatore tramite la classe NaiveBayes. Naive Bayes è un algoritmo di apprendimento automatico supervisionato che si basa

sul teorema di Bayes. È ampiamente utilizzato per la classificazione di testi e problemi di apprendimento automatico in cui le variabili di input sono indipendenti tra loro. L'algoritmo calcola la probabilità che una determinata istanza appartenga a ciascuna classe e assegna l'istanza alla classe con la probabilità più alta.

```
NaiveBayes nb = new NaiveBayes();
    nb.buildClassifier(structure);
    Instance current;
    while ((current = loader.getNextInstance(structure)) != null)
        nb.updateClassifier(current);
```

Quarto step: carica un file ARFF di nome "Test.arff" utilizzando la classe DataSource di Weka. Il DataSource fornisce un'interfaccia per accedere ai dati dal file ARFF. Successivamente, crea un oggetto Instances che rappresenta l'insieme di istanze del set di dati di test. Il metodo getDataSet() legge il file ARFF specificato e restituisce un oggetto Instances che rappresenta l'insieme di istanze. Infine, viene impostato l'indice della classe nell'insieme di istanze per consentire la classificazione corretta delle istanze di test. numAttributes() restituisce il numero totale di attributi nel set di dati e sottraendo 1, si ottiene l'indice dell'ultimo attributo, che di solito è l'attributo di classe. L'istruzione setClassIndex() imposta l'indice della classe per l'insieme di istanze in modo che l'algoritmo sappia quale attributo è la classe da predire.

```
DataSource test = new DataSource("Test.arff");
Instances teststructure = test.getDataSet();
teststructure.setClassIndex(teststructure.numAttributes()-1);
```

Quinto step: utilizza un classificatore Naive Bayes per classificare l'istanza di test e aggiorna il valore dell'attributo di classe dell'istanza con la classe predetta. In questo modo, è possibile utilizzare il classificatore addestrato per predire la

classe delle istanze di test e ottenere il valore predetto per l'attributo di classe. Viene chiamato il metodo `classifyInstance()` sul classificatore `nb` per classificare l'istanza di test specificata. `teststructure.instance(0)` restituisce la prima istanza dell'insieme di istanze del set di dati di test (`teststructure`). Il metodo `classifyInstance()` restituisce un valore `double` che rappresenta l'etichetta o la classe predetta per l'istanza di test. Viene richiamato il metodo `setClassValue()` sull'istanza di test per impostare il valore dell'attributo di classe con la classe predetta (`label`). `teststructure.instance(0)` restituisce l'istanza di test specificata (la prima istanza nel nostro caso) e `setClassValue()` imposta il valore dell'attributo di classe con il valore specificato (`label`).

```
double label = nb.classifyInstance(teststructure.instance(0));  
teststructure.instance(0).setClassValue(label);
```

Sesto step: il codice crea una nuova istanza di test utilizzando l'ultima istanza nel set di dati di test. Quindi utilizza un classificatore Naive Bayes per classificare l'istanza e ottiene la classe predetta. Infine, stampa il risultato della classificazione e assegna la classe predetta alla variabile `prediction`. Viene creata una nuova istanza (`newInst`) utilizzando l'ultima istanza dell'insieme di istanze del set di dati di test (`teststructure`). `teststructure.numInstances()` restituisce il numero totale di istanze nel set di dati di test, e sottraendo 1 si ottiene l'indice dell'ultima istanza. Viene chiamato il metodo `classifyInstance()` sul classificatore Naive Bayes (`nb`) per classificare la nuova istanza (`newInst`). Il metodo restituisce un valore `double` che rappresenta l'etichetta o la classe predetta per l'istanza di test. Viene stampato il risultato della classificazione. `structure.classAttribute().value((int) predNB)` restituisce la stringa corrispondente alla classe predetta per l'istanza (`predNB`). Quindi viene visualizzato un messaggio che indica quale classe è stata predetta dall'algoritmo. Il valore della classe predetta viene assegnato alla variabile `prediction`.

```
Instance newInst= teststructure.instance(teststructure.numInstances()-  
1);
```

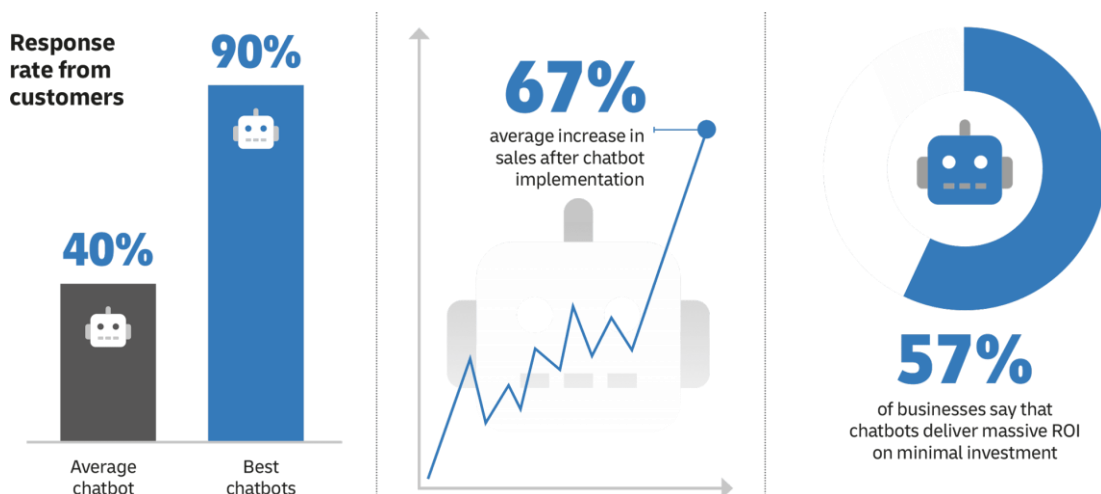
```
double predNB=nb.classifyInstance(newInst);  
System.out.println("L'AI dice che la figura é:  
"+structure.classAttribute().value((int) predNB)); //  
prediction=structure.classAttribute().value((int) predNB);
```

5. CONCLUSIONI

Riprendendo l'introduzione iniziale, abbiamo visto che l'intelligenza artificiale interattiva ha davvero il potenziale per rivoluzionare il modo in cui interagiamo con le macchine, rendendole più reattive e intuitive.

Poiché questo trend coinvolge algoritmi in grado di elaborare l'input umano e fornire una risposta adeguata, questa tecnologia ha diverse applicazioni, incluso il rilevamento facciale, i chatbot, gli assistenti digitali e molto altro. Nella logistica e nella supply chain, l'uso dell'IA interattiva migliora l'efficienza operativa e l'esperienza del cliente.

Aziende come DHL, ad esempio, prospettano ulteriori sviluppi in questo settore nel corso dei prossimi 5 anni, i quali avranno un impatto rilevante sulle strategie aziendali e promuoveranno una maggiore collaborazione tra persone e macchine. Ciò è attribuibile all'osservazione dei benefici e dell'aumento dei profitti derivanti dalle loro analisi.



I chatbot sono strumenti sempre più utili per gestire l'interazione con i clienti e semplificare le operazioni in una varietà di settori, inclusa la logistica. Possono ottimizzare la gestione e il controllo delle scorte, il tracciamento delle merci e gli orari di consegna, nonché l'aggiornamento del customer relationship

management (CRM) e del warehouse management system (WMS). L'adozione della tecnologia di IA interattiva all'interno della supply chain automatizza i flussi di lavoro e la gestione degli ordini, liberando i dipendenti delle operazioni per concentrarsi su compiti più complessi e ad alto valore aggiunto. Nel caso in cui un chatbot non sia in grado di completare una richiesta, essa viene assegnata a un operatore umano per ulteriori azioni, mentre il richiedente viene informato dello stato della richiesta.

Nel contesto del nostro programma per riconoscere le figure, abbiamo osservato un coinvolgimento attivo nel processo di apprendimento. Tuttavia, è importante notare che tale coinvolgimento può comportare un aumento del rischio di errore umano. Se commettiamo un errore, la macchina seguirà conseguentemente tale errore. E se questo viene ripetuto numerose volte, il rischio di avere un comportamento errato persistente da parte della macchina aumenta in modo significativo.

Risulta quindi fondamentale adottare misure adeguate per mitigare il rischio di errori umani ricorrenti nel processo di apprendimento del nostro programma di riconoscimento delle figure. Una strategia efficace consisterebbe nell'implementare meccanismi di controllo e validazione per verificare l'accuratezza e l'attendibilità delle informazioni fornite dagli umani durante il processo di addestramento del modello. Ciò potrebbe includere revisioni e verifiche incrociate da parte di esperti umani qualificati, nonché l'utilizzo di dataset di addestramento diversificati e rappresentativi per ridurre la possibilità di bias o errori sistematici. Inoltre, è essenziale monitorare costantemente le prestazioni del modello e intervenire tempestivamente per correggere eventuali deviazioni o comportamenti errati che potrebbero emergere. Con un approccio oculato e un continuo miglioramento del sistema, possiamo ridurre al minimo il rischio di errori umani e garantire un funzionamento affidabile e accurato del nostro programma di riconoscimento delle figure.

6. BIBLIOGRAFIA E SITOGRAFIA

1. **IBM: cos'è il Machine Learning :** [Online]
<https://www.ibm.com/topics/machine-learning>
2. **Microsoft – Introduzione al ML e ai suoi modelli** [Online]
<https://learn.microsoft.com/it-it/training/modules/introduction-to-machine-learning/2-what-are-ml-models>
3. **RedGate - Introduction to artificial intelligence, Kumar Abhishek,** [Online] 10 May 2022 <https://www.red-gate.com/simple-talk/development/data-science-development/introduction-to-artificial-intelligence/>
4. **J. Estevez, G. Garate and M. Graña, "Gentle Introduction to Artificial Intelligence for High-School Students Using Scratch,"** in IEEE Access, vol. 7, pp. 179027-179036, 2019, doi: [Online] 10.1109/ACCESS.2019.2956136.
<https://ieeexplore.ieee.org/document/8915693?denied=>
5. **Interactive Human Centered Artificial Intelligence: A Definition and Research Challenges** - Albrecht Schmidt - Ludwig-Maximilians-Universität München
56–61-Shneiderman-Human-Centered-AI-Winter-2021
6. [Online] <https://onlinelibrary.wiley.com/doi/epdf/10.1002/hbe2.117>
Mark Riedl
7. [Online] <https://www.cognizant.com/us/en/glossary/human-centered-ai>
8. **MIT: Machine learning, explained:** [Online]
<https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained>

9. <https://www.semanticscholar.org/paper/Interactive-Artificial-Intelligence:-Designing-for-Wenskovitch-North/4e7997058cc3054f219ba87bc4cc6ef16d00111b>
10. [Online]<https://www.imperial.ac.uk/jobs/description/ENG02216/research-assistant-associate-human-loop-machine-learning-systems-research>
11. [Online]<https://www.futurelearn.com/info/courses/advanced-data-mining-with-weka/0/steps/29464>
- 12. MNIST database - Wikipedia[Online]**
13. Documentation - Weka Wiki (waikato.github.io)[Online]
14. [Online]machine learning - Lightweight incremental classification of 1 dimensional data in java - Stack Overflow
- 15. Foundations of Machine Learning di Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar**
16. Shneiderman, Ben (2020), Human-centered artificial intelligence: Reliable, safe & trustworthy, International Journal of Human--Computer Interaction, Vol. 36, No. 6, p. 495-504.
17. Wenskovitch, John and North, Chris (2020), Interactive Artificial Intelligence: Designing for the "Two Black Boxes" Problem, Computer, Vol. 53, No. 8, IEEE, p.29-39 Mark O. Riedl (2019), Human-Centered Artificial Intelligence and Machine Learning, Human Behavior and Emerging Technologies , Vol. 1, No. 1, Wiley Online Library, p. 33-36.
18. Mhlanga, David (2022), Human-Centered Artificial Intelligence: The Superlative Approach to Achieve Sustainable Development Goals in the Fourth Industrial Revolution, Sustainability, Vol. 14, No. 13, MDPI, p. 7804

19. [Online] **DHL, Interactive AI** <https://www.dhl.com/global-en/home/insights-and-innovation/thought-leadership/trend-reports/ai-logistics.html#:~:text=The%20trend%20of%20Interactive%20AI,and%20provide%20a%20reasonable%20response.>