



UNIVERSITÀ  
DI SIENA  
1240

**DIPARTIMENTO DI ECONOMIA POLITICA E STATISTICA  
SCUOLA DI ECONOMIA E MANAGEMENT**

CORSO DI LAUREA IN SCIENZE STATISTICHE PER LE  
INDAGINI CAMPIONARIE

La statistica e la misurazione dell’“empatia”  
(digitale): un esperimento all’interno di una  
grande banca italiana

Relatore  
Prof. Stefano Duccio Gazzei

Tesi di laurea di  
Emanuele di Pietro

## Indice

<i>Abstract</i> .....	<i>pag.</i> 3
<i>Introduzione</i> .....	5
<b>Capitolo I:</b> .....	<b>6</b>
<b>1.1 Cos’è l’“empatia”?</b> .....	<b>6</b>
<b>1.2 Perché si studiano le emozioni?</b> .....	<b>7</b>
<b>1.3 Empatia, Emozioni ed Advocacy</b> .....	<b>9</b>
<b>1.4 La misura dell’Empatia</b> .....	<b>11</b>
<b>1.5 Valutazione attraverso l’NPS</b> .....	<b>14</b>
<b>1.6 Appraisal</b> .....	<b>20</b>
<b>1.6.1 Atteggiamento (Attitude)</b> .....	<b>21</b>
<b>1.6.2 L’uso dell’Appraisal</b> .....	<b>22</b>
<b>1.7 Relazione tra Appraisal e Analisi NPS</b> .....	<b>24</b>
<b>1.8 L’importanza delle parole</b> .....	<b>27</b>
<b>Capitolo II:</b> .....	<b>30</b>
<b>2.1 La banca: Storia</b> .....	<b>30</b>
<b>2.2 I dati</b> .....	<b>31</b>
<b>2.2.1 Qualità del dato</b> .....	<b>36</b>
<b>2.3 Analisi esplorativa univariata</b> .....	<b>37</b>
<b>2.3.1 Le Analisi del Net Promoter Score (NPS)</b> .....	<b>42</b>
<b>2.3.2 Cos’è successo durante il periodo Covid-19?</b> .....	<b>49</b>
<b>2.3.3 Le “colpe” della Banca</b> .....	<b>52</b>

<b>2.4 Come hanno vissuto il periodo i Clienti? .....</b>	<b>61</b>
<b>2.5 La felicità incide sulla raccomandabilità? .....</b>	<b>82</b>
<b>2.6 Un indicatore di attualità .....</b>	<b>88</b>
<b>2.7 È davvero possibile prevedere un'emozione?.....</b>	<b>103</b>
<b><i>Conclusioni</i> .....</b>	<b>122</b>
<b><i>Appendice A: Codici Python</i> .....</b>	<b>125</b>
<b><i>Bibliografia</i> .....</b>	<b>142</b>
<b><i>Sitografia</i> .....</b>	<b>144</b>

## Abstract

Lo scopo di questo elaborato è di andare a studiare a posteriori, prendendo in esame il primo periodo *Covid-19* che va da marzo a dicembre 2020, i possibili effetti che possono esserci stati sul rapporto banca-cliente. L'idea è che un periodo così strano, e che nessuno mai si poteva aspettare, sia stato una sorta di laboratorio in cui qualcosa sicuramente è successo dal punto di vista emozionale.

Nello specifico si vuole andare alla ricerca di un qualsiasi *framework*<sup>1</sup> (se esiste, oppure no), sempre parlando di emozioni e comportamenti, che in qualche modo abbia condizionato la relazione tra la banca e i suoi clienti.

Il punto di partenza è stata un'analisi *NPS* perché in un articolo<sup>2</sup> di alcuni studiosi è risultato fuori che in tale analisi, oltre ad esserci informazioni sulla ‘*valutazione delle performance*’ e sulla ‘*valutazione etica*’, possono essere trovate anche informazioni su una ‘*valutazione del sentimento*’, ossia lo stato d'animo di una persona.

Per poter estrarre questa informazione e riuscire ad avere una emozione, provata dal cliente al momento dell'intervista, è stata utilizzata una tecnica chiamata “*Appraisal*”<sup>3</sup>, la quale assegna una emozione ad un testo, messaggio, commento, o qualsiasi frase, in base alle parole che contiene. L'estrapolazione dell'emozione è stata fatta con il

---

<sup>1</sup> Framework (eng.) può essere tradotto come struttura o quadro strutturale

<sup>2</sup> LEWIS, C., & MEHMET, M. (2020). «*Does the NPS® reflect consumer sentiment? A qualitative examination of the NPS using a sentiment analysis approach*», Internatiol Journal of Market Research.

<sup>3</sup> MARTIN, J. R., & WHITE, P. R. R. (2004). «*The language of evaluation: Appraisal in English*», Berlin, Germany: Springer, «*Strategy Maps*», Harvard Business School Press.

*software Python*<sup>4</sup>, prendendo un programma su cui erano già state fatte varie prove in inglese e adattandolo, ovviamente, alle complessità della lingua italiana.

Dopo aver estrapolato con successo le emozioni è stata fatta un'analisi descrittiva del fenomeno. Successivamente sono stati notati i *patterns*<sup>5</sup> probabilmente più importanti, che mettevano in relazione le emozioni (*moods*) con i temi della *detrazione/promozione* e, in particolar modo, della *fedeltà*, ossia la volontà di un cliente di voler continuare l'esperienza presso la banca o di voler chiudere i rapporti. Per questo motivo è stata fatta una semplice domanda: “è possibile modellizzare tutto questo per avere una previsione del *mood*?”

Un primo tentativo, ritornando al discorso *Covid-19*, è stato fatto tenendo conto sia dello *spazio* (regione di un cliente) sia del *tempo* (data dell'intervista), con la costruzione di un indicatore, chiamato “indicatore di restrizione”, che associa ad ogni cliente il colore della propria regione al momento dell'intervista. Infine, è stato implementato il modello, il quale, ricostruendo le condizioni avute nel periodo marzo-dicembre 2020, potrebbe essere in grado di prevedere, in modo corretto, l'emozione di un cliente nel 62.7% dei casi. Questa previsione potrebbe essere poi utilizzata dalla banca per poter adottare un approccio diversificato per ogni emozione, con particolare attenzione per coloro che intendono chiudere i rapporti.

---

<sup>4</sup> È un linguaggio di programmazione

<sup>5</sup> *Pattern* (eng.) può essere tradotto con schema

## Introduzione

Le emozioni sono uno strumento con cui una persona interagisce con un'altra, fondamentali per il rapporto sociale.

Ovviamente si suddividono in positive e negative, ed ognuna appare a seconda della situazione in cui ci si trova.

La sfera delle emozioni non è limitata allo studio psicologico delle persone, oggi ci sono sempre più studi che legano le neuroscienze al mondo economico, il cosiddetto “neuromarketing”, una disciplina che va a studiare come il cervello umano si comporta in relazione a prodotti, servizi, pubblicità, etc.

Lo scopo è, ovviamente, di cercare di capire il comportamento della mente ai fini di strategie di mercato.

Inoltre, in un anno come il 2020 (primo anno Covid-19), il cervello delle persone è stato fortemente stressato dal mix di emozioni che si è creato a causa del connubio di situazioni estremamente difficili (lockdown e pandemia su tutte) con situazioni positive.

Da tali considerazioni nasce lo studio di questo elaborato, avendo come caso studio quello di una grande banca italiana, cercando di capire e rendere meno intricato il rapporto cliente-banca dal punto di vista emozionale, soprattutto in un periodo particolarmente strano come l'anno 2020, in cui la routine di ognuno è stata stravolta.

## Capitolo I:

### 1.1 Cos'è l'"empatia"?

La vita quotidiana delle persone è composta da tanti stati d'animo, tanti sentimenti che si mescolano tra loro a seconda del contesto (si prova felicità quando si riceve un aumento dello stipendio, tristezza quando si ha una giornata dura e difficile).

L'empatia è quella capacità che permette di immedesimarsi, mettersi nei panni di un altro individuo, che sia una persona, un animale o un personaggio finto, per capirne, interpretarne e condividerne pensieri e sentimenti. Diventa importante per stabilire un rapporto ed entrare, quindi, più facilmente in sintonia con chi si sta interagendo.

L'empatia aiuta a collaborare con altri, costruire amicizie, prendere delle decisioni morali e intervenire quando vediamo qualcuno vittima di bullismo.

Gli esseri umani iniziano a mostrare segni di empatia già fin dalla nascita; quindi, non è una capacità che si acquista con il tempo ma, come sostiene Giacomo Rizzolatti<sup>6</sup>, fa parte del corredo genetico della specie.

Usiamo tale parola una miriade di volte nella nostra vita quotidiana dando quasi per scontato il suo significato e il suo *modus operandi*<sup>7</sup>, in realtà però è decisamente più complesso di quello che sembra. Prima di tutto bisogna precisare che l'empatia non è banalmente un concetto prettamente di natura psicologica, ma è stato studiato e scoperto negli anni '90 da Giacomo Rizzolatti (e il suo *team*) che al comando di ciò ci sono i

---

<sup>6</sup> RIZZOLATTI, G., & SINIGAGLIA, C. (2006). «So quel che fai. Il cervello che agisce e i neuroni specchio», Milano, Raffaello Cortina Editore.

<sup>7</sup> Locuzione latina: "modo di operare"

neuroni specchio<sup>8</sup>, ossia una classe di neuroni che si attivano quando si svolge un’azione (utilizzo di una mano, della bocca, etc) o quando osserviamo un’altra persona svolgerla.

Grazie a questi neuroni specchio l’azione compiuta da altri si genera dentro noi stessi consentendo di “conoscere” l’altra persona.

Questo è esattamente ciò che accade quando si guarda, per esempio, un film drammatico in cui lo spettatore si immedesima nel personaggio del film, immaginando quello che proverebbe se si trovasse in una situazione simile a quella del film (una morte, la fine di una relazione, etc). Tutto ciò spesso risulta automatico e spontaneo.

## **1.2 Perché si studiano le emozioni?**

Nel paragrafo precedente è stato accennato che l’empatia fa parte delle persone e incide nelle scelte che quest’ultime fanno quotidianamente (al supermercato, al lavoro, con gli amici, etc.). Con la globalizzazione poi, che ha mutato notevolmente le interazioni (anche incrementandole), l’emotività ha preso ancora più piede, diventando ancora più significativa.

Un focus va fatto sul mondo del lavoro. Fino a qualche anno fa si andava sul posto di lavoro con una eccessiva serietà, lasciando all’esterno ogni tipo di emozione e questo andava ad impattare in modo negativo sui dipendenti di ogni grado, fino ad arrivare al

---

<sup>8</sup> PELLEGRINO G., FADIGA L., FOGASSI L., GALLESE V., & RIZZOLATTI G. (1992). «*Understanding motor events: a neurophysiological study*». Exp. Brain Res. 91: 176-180.

*CEO*<sup>9</sup>.

Questo perché dopotutto siamo esseri umani e non può essere rinnegata la nostra natura, la nostra emotività, che potrebbe portare solamente a situazioni di stress fisico e psicofisico.

Da qui l'introduzione delle emozioni nell'ambito aziendale, poiché creando un clima meno serio e pesante, con la presenza anche di psicologi per i *team* delle risorse umane per capire al meglio le sensazioni della forza lavoro, si hanno impatti positivi sulle *performance*<sup>10</sup>.

Di fondamentale importanza è ovviamente, il cliente, dato che ormai le esigenze e gli interessi di una qualsiasi azienda sono basati su di esso, su ciò che costui ha bisogno o di cui avrà bisogno, magari, in futuro.

Nell'ottica moderna di azienda il cliente svolge un ruolo di particolare importanza per la sopravvivenza della stessa, ruolo che può essere definito imprescindibile, poiché se fino a qualche anno fa le aziende avevano il solo scopo di vendere il più possibile, oggi questo è cambiato.

Si passa da avere l'azienda al centro di tutto, all'avere il cliente, perché non si offre più un prodotto a quest'ultimo, ma si interagisce per rispondere a delle semplici e banali domande del tipo:

-A chi sarà destinato questo prodotto/servizio?

---

<sup>9</sup> CEO (*Chief Executive Officer*) è il corrispettivo americano di *Amministratore Delegato (AD)*

<sup>10</sup> *Performance* (eng.): prestazione

-Lo troverà utile e/o interessante?

-Il prodotto/servizio sarà all'altezza delle sue aspettative?

-Che esigenze ha questo cliente?

Si inizia, quindi, ad instaurare un vero e proprio rapporto con i clienti proprio grazie al meccanismo dell'empatia e al suo calarsi nei panni altrui.

In questo modo il cliente si sentirà “ascoltato”, si fiderà dell’azienda e penserà che i suoi bisogni verranno soddisfatti.

### **1.3 Empatia, Emozioni ed Advocacy**

Con il termine “*Advocacy*” si intende una filosofia aziendale di gestione della relazione con il cliente basata proprio sulla comprensione dei suoi bisogni, soprattutto sotto forma di esperienza emozionale (*customer emotional journey*), per mettere in atto un modello di servizio capace di allinearsi, oltre che alle aspettative in termini di prodotti e servizi, anche al suo stato d’animo.

Ma perché c’è bisogno di fare questo? Perché è interessante e utile capire a fondo ciò che un cliente pensa di noi, del nostro modo di relazionarci con lui, di vendere prodotti o di promuovere il nostro *brand*?

L’ *Advocacy* diviene importante poiché è quel *framework* di attività, di funzioni aziendali che consentirà di mettere il cliente al centro di tutto, di renderlo il vero protagonista del mercato e di promuovere ciò che lui pensa dell’azienda stessa, sia le

sensazioni positive che quelle negative.

Un occhio di riguardo, ovviamente, va sulle sensazioni positive perché sono quelle che fanno bene alla salute dell'azienda e possono essere considerate come una pubblicità gratuita che il cliente fa nei confronti dell'azienda che è stata in grado di ascoltarlo. Gratuita perché viene fatta dal cliente stesso e quindi l'azienda non dovrà sostenere (apparentemente) nessun costo, ma, nonostante ciò, risulta molto più potente di qualsiasi attività pubblicitaria che l'azienda possa intraprendere, questo perché la recensione positiva di una persona, che magari si conosce, fa più effetto di una pubblicità, nonostante in quest'ultima spesso vengano citati solo i pregi di un determinato prodotto/servizio.

Per questo motivo l'*Advocacy* può essere considerata come l'"*avvocato*" del cliente.

L'empatia, però, non entra in gioco solamente quando bisogna instaurare questo rapporto tra cliente e azienda, fa molto di più perché permette, oltre che fare una distinzione tra i vari clienti riguardo caratteristiche come reddito, età, interessi, livello di istruzione, di considerare anche il campo delle emozioni, per poter approcciare in maniera diversificata ogni tipologia di cliente.

Questo nuovo modo di vedere le cose diventa molto interessante, non si studiano più i clienti solamente riguardo le loro variabili sociodemografiche, ma vi è un concetto ancora più semplice, ossia il proprio stato d'animo attuale. Può sembrare strano, o forse no, ma ogni volta che una persona interagisce con un'altra persona, un animale, un

oggetto, viene inondata da una marea di emozioni che sono importanti perché è grazie a loro che si prendono le decisioni, si effettuano scelte e per questo bisogna stare anche attenti, per non rischiare di prendere decisioni sbagliate.

Ovviamente tutto ciò viene applicato in un'ottica aziendale, soprattutto in campo *marketing* e ciò che questo elaborato vuole studiare è proprio la differenza delle reazioni che un cliente può avere a causa di una determinata emozione; per esempio, una persona che ha appena subito un lutto (triste) si comporterà in modo del tutto differente rispetto ad una persona che ha appena vinto alla lotteria (gioia), seppur riguardo le stesse interazioni con l'azienda, che nel caso di una banca, che è proprio l'istituto oggetto di studio di questo elaborato, possono essere un prestito, un mutuo, un prelievo, etc.

## 1.4 La misura dell'Empatia

Una volta stabilito quanto sia importante per il successo di un'azienda studiare i propri clienti e scoprirne gli stati d'animo, nasce un ulteriore problema, ossia "Come misurare questi stati d'animo?"

Per poter fare ciò è possibile utilizzare in questo caso uno strumento molto potente e anche abbastanza recente, poiché nato solamente negli ultimi 20 anni, ossia il *Net Promoter Score (NPS)*.

L'NPS nasce nel 2003 da Fred Reichheld che lo introdusse nell' *Harvard Business*

*Review*<sup>11</sup> come strumento per prevedere la crescita delle vendite future data la sua semplicità per implementare, analizzare e comunicare. Oggi, invece, il suo scopo si è spostato sul misurare lo stato di salute di un *brand/azienda* come vero e proprio indicatore di *performance* dopo che la letteratura ha smentito, o comunque non confermato, la relazione tra il punteggio NPS e la crescita di un'azienda<sup>12</sup>.

Nasce dall'esigenza di avere uno strumento molto semplice, e facilmente interpretabile da chiunque, per determinare una valutazione, o meglio un punteggio, della soddisfazione del cliente riguardo l'azienda/prodotto/servizio.

L'NPS misura il passaparola come proxy della soddisfazione del cliente ed è ben noto che il passaparola fornisce un mezzo di pubblicità non retribuita e può influenzare l'immagine del marchio.

---

<sup>11</sup> REICHHOLD, F. F. (2003). «*The one number you need to grow*». Harvard Business Review, 81, 46–55.

<sup>12</sup> BAEHRE, S., O'DWYER, M., O'MALLEY, L., & LEE, N. (2021). «*The use of Net Promoter Score (NPS) to predict sales growth: insights from an empirical investigation*», Journal of the Academy of Marketing Science.

Si basa tutto su una semplice ma fondamentale domanda:

-Quanto raccomanderesti l'azienda/prodotto/servizio ad un amico/parente da 0 a 10?

Oltre a questa domanda, la cui risposta banalmente è un numero compreso tra 0 = *per niente* e 10 = *del tutto*, vi è un ulteriore domanda, in questo caso aperta, dove si chiede al cliente di scrivere una motivazione (in gergo *Verbatim*) sul perché del voto dato, per cogliere al meglio tutte le sfaccettature e i motivi.

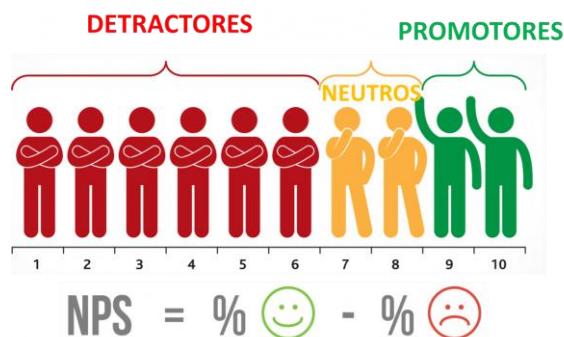
Le votazioni da 0 a 10 vengono successivamente suddivise in 3 fasce nel seguente modo:

-*Promoters* (o Promotori): punteggio da 9 a 10;

-*Neutros* (o Neutrale): punteggio da 7 a 8;

-*Detractors* (o Detrattori): punteggio da 0 a 6.

*Figura 1. Net Promoter Score (NPS)*



*Fonte: Health Marketing*

Per calcolare l'NPS non bisogna fare altro che la differenza tra %Promoters e

*%Detractors*, ottenendo così un valore compreso tra -100 e +100 dove ovviamente un valore positivo sottolinea il fatto che ci siano più *Promoters* rispetto ai *Detractors* e che, in sintesi, l'azienda sta procedendo bene.

Il *Verbatim* è una domanda aperta che quindi lascia spazio alla fantasia del cliente, il quale può scrivere qualsiasi cosa e più queste motivazioni saranno lunghe e dettagliate, specificando per bene i punti di forza e le debolezze, e più saranno precise e utili le informazioni che il *Data Scientist*<sup>13</sup> potrà ricavare.

## 1.5 Valutazione attraverso l'NPS

Per effettuare un'analisi di valutazione attraverso il *Net Promoters Score* bisogna iniziare dal questionario, il quale verrà accuratamente preparato e somministrato ad un campione rappresentativo della clientela totale; in questo caso il campione è di circa 45.000<sup>14</sup>.

Una volta somministrato il questionario e ricevute le risposte fornite dai clienti, si può passare allo studio di tali dati che saranno poi presi in considerazione dall'azienda stessa, in primis, come strumento di soddisfazione totale della clientela, ma anche come mezzo per indirizzare al meglio le strategie aziendali.

---

<sup>13</sup> Esperta figura professionale che si occupa della *scienza dei dati* (*Data Science*)

<sup>14</sup> Approfondimento paragrafo 2.2.1

Recentemente sono stati pubblicati alcuni studi<sup>15</sup> basati sull'indicatore NPS dove è stato dimostrato che la risposta fornita da un cliente, direttamente o indirettamente, non “copre” un unico aspetto, ma può essere investigata per capire il pensiero del cliente su almeno tre tematiche alquanto differenti tra loro:

1. Valutazione delle *performance*;
2. Valutazione etica;
3. Valutazione del “*sentiment*”.

Si hanno tre tipologie di valutazioni che possono essere sviluppate grazie ai dati ricavati dal questionario NPS che ricoprono tre sfere diverse.

Perché si studiano questi tre aspetti?

1. La valutazione delle *performance* nasce dall'esigenza di un'azienda di monitorare costantemente la situazione “clienti”, di capire come questi valutano i vari brand, i vari prodotti e/o servizi offerti. È una valutazione incentrata totalmente sul voto che i clienti assegnano all'azienda, ossia sul grado di “raccomandabilità” che danno in base alle esperienze avute in passato con la stessa. Il voto diventa sintesi dell'esperienza del cliente, all'apparenza può sembrare un semplice numero compreso tra 0 e 10, ma in realtà racchiude ben altro poiché permette di capire se il contatto azienda-cliente ha avuto un

---

<sup>15</sup> LEWIS, C., & MEHMET, M. (2020). «Does the NPS® reflect consumer sentiment? A qualitative examination of the NPS using a sentiment analysis approach», Internatiol Journal of Market Research.

risposto positivo o negativo, proprio ciò che interessa ad un'azienda per capire come muoversi.

La valutazione non è però così banale, calcolando inizialmente l'NPS come  $\%Promoters - \%Detractors$  si avrà già una buona idea dell'andamento, ma non basta poiché si rimane sul generale; un'azienda che ha un NPS pari a +6, un valore positivo che tutto sommato indica un buon stato di salute è in realtà così?

Un'azienda è un'entità organizzata in più sfere che a loro volta si occupano di compiti diversi dove tutti, ovviamente, sono accomunati da uno scopo comune, il benessere del cliente e, quindi, dell'azienda stessa; per questo motivo valutare in generale non serve a molto dato che una sfera aziendale può aver fatto un ottimo lavoro, mentre un'altra decisamente no.

Per avere delle idee un po' più chiare sulla performance dell'azienda bisogna tener conto anche del *Verbatim*, cioè di tutte quelle motivazioni, di quei commenti che i clienti hanno espresso in maniera del tutto libera per spiegare il voto dato.

Su questi *Verbatim* viene fatta un'analisi testuale, in gergo *Text Mining*, da cui è possibile ricavare le varie sfere, chiamate *drivers* che rappresentano i fattori critici su cui l'azienda deve stare attenta e investire in qualche modo.

I fattori critici vengono “raccontati” proprio dai clienti, perché sono ricavati dalle parole che questi utilizzano nei *Verbatim*; inoltre, bisogna precisare che

questi fattori critici (es. prezzo di un prodotto/servizio, competenze del personale, immagine dell'azienda, sito web, etc.) altro non sono che i punti di forza e di debolezza che i clienti sottolineano.

Ovviamente eccellere in tutto è pura utopia ed infatti non è così fondamentale quanto sembra, ma è fondamentale eccellere in quei punti critici in cui il cliente magari si aspetta di più e di cui l'azienda ne ha fatto una vera e propria filosofia di vita; *Apple* propone *smartphone* di alta qualità a prezzi molto elevati (attenzione sul *driver delle performance* del prodotto, ma non su quello del prezzo), mentre *Ikea* propone mobili a prezzi ridotti ma che, magari, ne risentono in termini di comodità (attenzione sul *driver del prezzo* ma non su quello delle *performance*).

Per ogni *driver*, quindi, viene calcolato un *NPS Coin* che indica se quel determinato fattore critico sta *Creando Valore* (es. *driver* del prezzo per *Ikea*) o *Distruggendo Valore* (es. *driver* del prezzo per *Apple*) e in base alla “filosofia” adottata dall'azienda si potrà capire su quale *driver* bisogna prendere delle decisioni per migliorare e su quali continuare così perché magari si sta facendo già molto bene.

2. La valutazione etica è strettamente collegata all’ “Etica aziendale” che oggi, più che mai, grazie a tematiche sensibili sulla cresta dell’onda come la sostenibilità e l’ambiente, è diventata di fondamentale importanza. Alle aziende non basta più

solo sapere di fare un buon lavoro, di soddisfare egregiamente i propri clienti e di avere delle buone entrate, ma è diventato importante ormai essere coerenti con gli obiettivi posti inizialmente. Per capirci, un'azienda può aver fatto un ottimo lavoro dal punto di vista delle *performance*, ma eticamente può aver fallito, svolgendo magari un lavoro che le persone reputano non idoneo, non necessario, o comunque non primario (es. sala scommesse).

È di particolare interesse, quindi, valutare anche se ciò che è stato fatto viene approvato dai propri clienti sotto vari punti di visti, anche quello umano.

Anche qui si lavora sui *Verbatim* perché, come visto in precedenza, racchiude tutto ciò che un cliente pensa dell'azienda, compreso se si sia comportata in modo etico o no.

### 3. Infine, si ha la valutazione del “*sentiment*”, ossia lo stato d'animo dei clienti.

Delle tre è sicuramente la più particolare, nel senso che solo ultimamente si è iniziato a fare studi sulla correlazione tra *sentiment* e il modo di comportarsi dei clienti nella realtà aziendale, per questo motivo risulta ancora più significativa una valutazione simile.

Importante è precisare che il *sentiment* può dipendere da più fattori, ma in sintesi se ne possono definire due che per noi sono il fulcro della nostra analisi:

- Condizioni esterne
- Banca

In entrambi i casi le motivazioni possono essere tra le più svariate, nel primo caso spesso vi è lo stress, problemi economici, problemi di salute, etc. Nel secondo, invece, possiamo trovare tutto ciò che riguarda la banca come un personale non educato, un modo di esprimersi non idoneo al tipo di persona che si ha di fronte e tanto altro.

Diventa fondamentale, quindi, studiare caso per caso e stabilire se il *sentiment* è dovuto da condizioni esterne o dalla banca e, ovviamente, prendere in considerazione soprattutto quest'ultimo fattore.

A livello operativo, l'NPS non fornisce indicazioni strategiche sui miglioramenti che potrebbero essere fatti, sulle percezioni e/o sulle azioni dei concorrenti o su quanti possano essere i nuovi clienti acquisiti perché sono compiti che spettano al *management*.

## 1.6 *Appraisal*

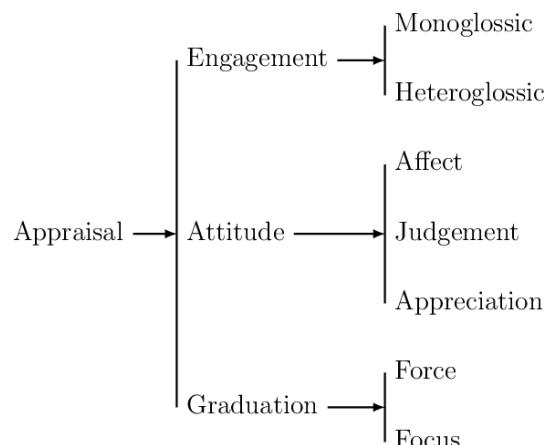
La tecnica utilizzata per estrarre il sentimento delle persone è di recente sviluppo ed è chiamata *Appraisal*<sup>16</sup>, ed è qui lo strumento *core*; inoltre, anche se gli approcci automatizzati che utilizzano frasi e parole chiave si basano su dati di massa, non sempre riescono a identificare tutte le sfumature associate.

Il sistema dell'*Appraisal*, che si basa sugli strati semanticci di un discorso, fu proposto da J. R. Martin e P. White.

L'*Appraisal framework* è strutturata su tre principali sistemi semanticci (o domini) che interagiscono tra di loro: *Atteggiamento (Attitude)*, *Engagement* e *Graduation*.

Il più interessante per questo elaborato è sicuramente l'*Atteggiamento*.

Figura 2: Schema "Appraisal"



Fonte: "The Language of Evaluation: Appraisal in English", J.R. Martin, P. R. R. White

---

<sup>16</sup> MARTIN, J. R., & WHITE, P. R. R. (2004). «*The language of evaluation: Appraisal in English*», Berlin, Germany: Springer, «*Strategy Maps*», Harvard Business School Press.

### 1.6.1 Atteggiamento (*Attitude*)

Il modo con cui una persona sente qualcosa o qualcuno è chiamato atteggiamento, è un sistema che coinvolge tre regioni semantiche (*Stato d'animo*, *Giudizio* e *Apprezzamento*) che coprono rispettivamente ciò che viene chiamato emozione, etica ed estetica. L'emozione è, probabilmente, al centro di queste regioni poiché è la risorsa espressiva con cui nasciamo e incarniamo fisiologicamente al momento della nascita.

- ***Stato d'animo (Affect)***: si occupa di registrare le emozioni positive o negative espresse in un testo/messaggio come, ad esempio, felicità, tristezza, speranza e soddisfazione. Ci sentiamo felici o tristi? Fiduciosi o ansiosi?

Interessati o annoiati? Nel seguente esempio viene fuori un senso di tristezza, di malinconia, quando un membro della *Stolen Generation*<sup>17</sup> parla della separazione dai suoi fratelli.

[...Il **dolore** è arrivato per mia sorella minore e i miei due fratelli pensando che non li avrei più rivisti. Il giorno in cui ho lasciato l'orfanotrofio è stato **molto triste** per me. Ero **infelice**...]<sup>18</sup>

Diventa quindi importante l'uso delle parole, ognuna ha un proprio significato e ognuna (a differenza di altre) dice qualcosa, permette di orientare il cervello su una determinata azione, un determinato pensiero.

---

<sup>17</sup> *Stolen generation*, in italiano *generazione rubata* è il nome con cui vengono indicati quei bambini australiani aborigeni allontanati dalle proprie famiglie da parte dei governi federali australiani ai sensi di alcune norme parlamentari vigenti nei singoli Stati. Partendo dal 1869, gli allontanamenti forzati continuarono fino al 1969, anche se in alcuni luoghi proseguirono sino agli anni '70.

<sup>18</sup> AUSTRALIAN HUMAN RIGHTS COMMISSION (2 August 1995). «Bringing them Home - preliminary».

- **Giudizio (Judgment)**: relativo alla valutazione di moralità di una persona, veridicità e capacità, riguarda gli atteggiamenti nei confronti del comportamento come quando si ammira, o si critica, o si loda, etc. Esempio:
- **Apprezzamento (Appreciation)**: si riferisce alla valutazione di un qualcosa o qualcuno in base alle loro *performance* o qualità estetiche. Esempio:

[...Praticamente perfetto, con dettagli impeccabili, personaggi sorprendentemente originali e una trama avvincente...]<sup>19</sup>

Describe il valore di qualcosa. Martin e White sottolineano che questa classe dipende molto dal discorso e dai suoi partecipanti perché il presunto valore di una cosa è variabile da persona a persona.

### 1.6.2 L'uso dell'Appraisal

Nel paragrafo precedente è stata spiegata, in teoria, la tecnica dell'*Appraisal*, la quale si basa, in parole povere, sui termini contenuti in un testo, messaggio, una qualsiasi frase.

Prendendo come esempio una banale frase, “*sono felice*”, andandola a leggere risulta facile capire che in tale frase vi è un senso di felicità, di gioia, mentre se avessimo considerato “*non sono felice*”, il senso sarebbe diventato più cupo, di tristezza.

La tecnica è applicata attraverso un *text miner*<sup>20</sup>, in cui per ogni emozione è stata creata

---

<sup>19</sup> ONDAATJE, M. (2000). «Anil's Ghost», Toronto, Vintage.

una lista di parole (es. il termine “*felice*” porta positività, “*depresso*” porta negatività).

La lista è stata presa da progetti americani e, ovviamente per adattarla al caso italiano, tradotta.

Bisogna fare attenzione però, il concetto alla base dell’algoritmo di *text miner* è abbastanza semplice (associo un’emozione in base alle parole che trovo in una frase), ma nella pratica nascono spesso molti problemi; la lingua italiana, rispetto all’inglese, conta molti più particolarismi e forme delle parole.

Un esempio importante sono le negazioni, mettendo la parola “*felice*” in una frase del tipo “*io sono felice*”, il *text miner* ci dirà gioia in modo corretto, se invece mettiamo la stessa parola in una frase come “*io non sono felice*”, il *text miner* ci dirà comunque gioia perché troverà ugualmente la parola “*felice*”. In questo caso dovremmo aggiungere nella lista della tristezza “*non felice*”, così l’algoritmo<sup>21</sup> vedendo anche la negazione non sbaglia. Purtroppo, le casistiche sono tante (“*sono poco felice*”, “*sono abbastanza felice*”, “*potrei essere più felice*”, etc.) e non è possibile considerarle tutti in un primo elaborato e soprattutto con così poco tempo a disposizione; quindi, molto probabilmente ci saranno degli errori durante l’assegnazione.

---

<sup>20</sup> Deriva da *text mining*, tecnica che estrae informazioni utili da messaggi, testi, frasi.

<sup>21</sup> In Appendice è presente il codice del text miner

## 1.7 Relazione tra *Appraisal* e Analisi NPS

Che relazione può esserci tra una tecnica specializzata per la classificazione di testi, frasi, relativamente alle parole usate in ciascuno e un'analisi che si basa su ciò che il cliente racconta riguardo un'azienda per migliorarsi in futuro?

La risposta l'hanno data due scienziati, Clifford Lewis e Michal Mehmet, pubblicando un articolo molto importante sull'*International Journal of Market Research*, dal titolo “*Does the NPS® reflect consumer sentiment? A qualitative examination of the NPS using a sentiment analysis approach*”.

L'articolo parla proprio dell'uso dell'*Appraisal* in un contesto di analisi NPS, sostenendo che la suddivisione in *Stato d'animo*, *Giudizio* e *Apprezzamento* che viene proposta in quella tecnica, può essere trovata nell'analisi NPS.

Questo come?

L'analisi NPS, come è stato detto in precedenza, è composta tendenzialmente da due elementi: un voto (da 0 a 10) e un *Verbatim*, ossia un commento fatto dal cliente riguardo il voto dato (una recensione).

Applicando la tecnica dell'*Appraisal* su quanto detto, i due studiosi hanno notato e dimostrato che il voto NPS racchiude ciò che fa parte del Giudizio, così come il *Verbatim* racchiude ciò che si trova nell'*Apprezzamento* (commenti positivi e negativi).

Rimane fuori solamente lo *Stato d'animo*, che a primo impatto non corrisponde a nessuno dei due elementi racchiusi nell'analisi NPS, ma che può, dicono Lewis e

Mehmet, essere ricavato dal *Verbatim*, poiché essendo un commento dato dal cliente sulla sua esperienza con l'azienda riguardo prodotti/servizi, non ha al suo interno solamente una recensione (che sia positiva o negativa), ma anche ciò che il cliente prova a livello emozionale perché ogni essere umano è diverso dagli altri, ognuno reagisce a qualcuno/qualcosa in base a come si è formato il proprio cervello.

Questo articolo è fondamentale, perché per la prima volta ci si spinge così in là in questo mondo del *neuromarketing*<sup>22</sup> che racchiude anche l'umore del cliente.

Successivamente, la cosa importante diventa capire quanti e quali sono effettivamente questi “*moods*” che esprimono i clienti. Esistono vari modi, varie classificazioni fatte da più studiosi; c’è, per esempio, una divisione in sette emozioni principali, in sei ed anche in cinque, che probabilmente è la più famosa poichè apparsa nel noto film della Pixar *Inside Out*<sup>23</sup>.

Quella che verrà utilizzata in questa analisi è la suddivisione proposta da Paul Ekman<sup>24</sup>, uno psicologo statunitense.

In uno dei suoi esperimenti Ekman decise di girovagare intorno al mondo fotografando persone di ogni luogo per vedere se le espressioni facciali erano influenzate da cultura, tradizioni, posizione geografica.

---

<sup>22</sup> Recente disciplina che fonde il *marketing tradizionale* (economia) con *neurologia* (medicina) e *psicologia* (scienze comportamentali) per cercare di capire cosa accade nel cervello di una persona a contatto con prodotti, marche o pubblicità, con l’obiettivo di determinare le strategie che spingono all’acquisto.

<sup>23</sup> *Inside Out* è un film d’animazione del 2015, diretto da Pete Docter e Ronnie del Carmen; prodotto dai Pixar Animation Studios, in co-produzione con Walt Disney Pictures

<sup>24</sup> EKMAN, P. (1992). «Are There Basic Emotions?». Psychological Review, 99(3), 550-553.

Il risultato fu che in ogni parte del mondo, a partire da quelle più industrializzate e globalizzate fino a quelle più remote dove i mass-media ancora non erano arrivate, le espressioni erano sempre le stesse, che quindi una fotografia rappresentante una persona gioiosa poteva essere interpretata correttamente in qualsiasi angolo della terra.

Così, alla fine, definì sei emozioni principali:

1. Rabbia
2. Gioia
3. Sorpresa
4. Tristezza
5. Paura
6. Disgusto

Nel 2017, inoltre, uno studio della University of California ha proposto addirittura uno scenario che comprendeva 27 emozioni primarie<sup>25</sup>.

---

<sup>25</sup> COWEN, A. S., & KELTNER, D. (2017). «*Self-report captures 27 distinct categories of emotion bridged by continuous gradients*», PNAS, September 5, 2017.

## 1.8 L'importanza delle parole

Nei paragrafi precedenti è stato detto che le persone, durante il corso della vita o di una semplice giornata si trovano ad affrontare situazioni più o meno scomode e di conseguenza di comportarsi diversamente in ognuna di essa. Risulta facile pensare che un lutto porta tristezza così come la nascita di un figlio porta gioia, ma non è tutto così banale come può sembrare, da una parte dipende da ciò che accade quotidianamente, dall'altra dalle parole che vengono utilizzate.

A primo impatto il modo di parlare può sembrare una conseguenza della situazione circostante, ma non è così. Ci sono numerosi testi, articoli scientifici, che dimostrano come il modo di parlare di una persona incide direttamente sulla stessa (e anche su chi ascolta), in modo positivo o negativo, perché ogni singola parola attiva una specifica area del cervello.

Quando ripensiamo ad un momento felice il nostro umore migliora, quando ripensiamo ad un momento triste peggiora, perché tutto questo?

Il nostro cervello collega ogni parola ad un determinato contesto che, a sua volta, permette di rilasciare nel nostro corpo una moltitudine di sostanze chimiche (dopamina, serotonina, etc.), quindi potremmo considerare il nostro umore come un mix di queste sostanze chimiche che, a loro volta, dipendono dalle parole che quotidianamente vengono usate.

Nel suo libro “Basta Dirlo”<sup>26</sup>, Paolo Borzacchiello facendo riferimento a numerosi studi afferma che le parole che si pensano e pronunciano attivano nel cervello una ricerca semantica, la quale richiama una serie di idee collegate a quella parola.

Per esempio, scrivendo la parola “giraffa” automaticamente chi ascolta inizierà a pensare, ad avere una serie di idee su quella parola, collegandola ad un animale con il collo lungo, maculato etc. Non servirà specificare tutte queste caratteristiche dell’animale poiché sono comprese in quell’unica parola. Questo per far capire che una parola non è solamente una combinazione di sillabe ma molto più, contiene altre parole o addirittura immagini, suoni e come si diceva prima sostanze chimiche, ormoni.

Le parole fanno di più, sono in grado di far provare paura, una minaccia, di far restare una persona paralizzata per il forte spavento, aumentare il battito cardiaco, la pressione e tanto altro, il tutto in meno di  $\frac{1}{2}$  secondo per migliaia di volte al giorno. Da qui il perché del non pensare negativo (occhio però a non pensare troppo positivo) poiché rievoca nella persona sensazioni non piacevoli e il benessere dipende proprio dal tipo di chimica che si ha dentro.

Le sostanze chimiche che si sviluppano nel corpo e responsabili delle nostre “emozioni” sono varie:

- CORTISOLO: chiamato anche “ormone dello stress”;

---

<sup>26</sup> BORZACCHIELLO P. (agosto 2021). «*Basta dirlo-Le parole da scegliere e le parole da evitare per una vita felice*», Mondadori Libri S.p.A., Milano, I edizione.

- DOPAMINA: è la sostanza collegata alla promessa, all'aspettativa, ciò che ci fa stare bene;
- SEROTONINA: ormone della fiducia e della sicurezza in se stessi;
- ADRENALINA: ormone dell'eccitazione;
- OSSITOCINA: ormone dell'empatia e più in generale dell'affetto;
- TESTOSTERONE: ormone della sessualità, diminuisce i rischi di depressione;
- ENDORFINA: ormone della felicità;

## Capitolo II:

### 2.1 La banca: Storia

L'istituto oggetto d'analisi è uno dei più grandi gruppi bancari fondato oltre 100 anni fa.

È presente in molti paesi del mondo e può contare più di 150 mila collaboratori solamente in Europa dove è attiva su più mercati.

Promuove l'idea di sostenibilità, tema molto caldo oggi, rispettando tutte le varie forme di diversità, integrando soprattutto i giovani per valorizzare al meglio il futuro delle nuove generazioni restando al passo con gli sviluppi quotidiani.

Per quanto riguarda il territorio italiano, la banca è presente in ogni regione e si occupa del rilievo nel credito al consumo, *leasing, factoring, cash management*, pagamenti internazionali, servizi on line, microcredito, oltre che a servizi di *investment banking* e gestione dei rischi.

## 2.2 I dati

I dati, raccolti tramite questionario somministrato ai clienti, sono stati raccolti nel periodo che va dal 24 marzo 2020 al 31 dicembre 2020, periodo che ricopre proprio la prima Emergenza Covid, pandemia che, purtroppo, ha colpito, e sta colpendo, non solo il nostro Paese ma l'intero pianeta, mettendo in ginocchio tantissime realtà e soprattutto imprese, portando al fallimento numerose di esse.

Il dataset conta di 45713 osservazioni (clienti)<sup>27</sup>.

Questo lavoro di somministrazione del questionario oggi è diventato di vitale importanza poiché strumento principale per avere resoconti dai propri clienti, che ricordiamo essere ormai il centro di ogni attività imprenditoriale, sulla soddisfazione riguardo prodotti/servizi, eventuali problemi, consigli e tanto altro.

Le variabili che verranno trattate sono le seguenti:

- ID: è l'identificativo di ogni cliente che va da 1 a 45713;
- WAVE (ONDATA): composta da dieci ondate (una per mese), indica in quale ondata è stato effettuato il questionario dal cliente;
- PERSONAS: è una segmentazione effettuata sul cliente sulla base di tre

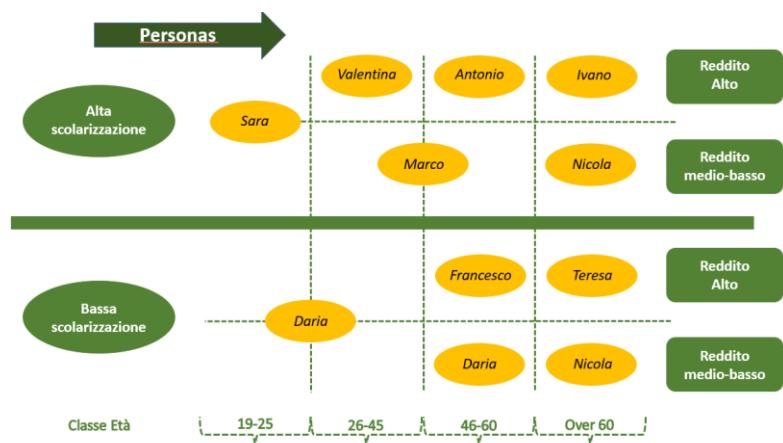
---

<sup>27</sup> Approfondimento paragrafo 2.2.1

variabili: ETÀ, REDDITO e SCOLARIZZAZIONE da cui si ricavano otto profili. I profili (ricavati da un'altra analisi fatta in precedenza da un'altra laureanda) sono i seguenti:

- 1) Antonio: 46-60 anni, alta scolarizzazione, reddito alto
- 2) Daria: 19-60 anni, bassa scolarizzazione, reddito medio-basso
- 3) Francesco: 46-60 anni, bassa scolarizzazione, reddito alto
- 4) Ivano: 60+, alta scolarizzazione, reddito alto
- 5) Marco: 26-60 anni, alta scolarizzazione, reddito medio-basso
- 6) Nicola: 60+, alta/bassa scolarizzazione, reddito medio-basso
- 7) Sara: 19-25 anni, alta scolarizzazione, reddito medio-alto
- 8) Teresa: 60+, bassa scolarizzazione, reddito alto
- 9) Valentina: 26-45, alta scolarizzazione, reddito alto.

*Figura 3. Segmentazione Personas*



*Fonte: Elaborazione personale, PowerPoint*

- NIELSEN: è anch'essa una segmentazione del cliente fatta sulla base di tre variabili che in questo caso sono: SPORTELLO (aver avuto contatti con lo sportello di una banca negli ultimi 12 mesi), GESTORE (aver avuto contatti con gestori, come consulenti) e INTERNET (aver avuto contatti con la banca in modo digitale);
  - 1) Fantasma: nessun tipo di contatto;
  - 2) Digito Relazionale: contatto su internet e con gestori;
  - 3) Multicanale: contatto di ogni tipo;
  - 4) Relazionale: contatto solo con gestori;
  - 5) Tradizionale: contatto con sportello e gestori;
  - 6) Digitale Puro: solo contatti internet;
  - 7) Digito Transazionale: contatti internet e con sportello;
  - 8) Transazionale Sportello: solo contatti con sportello;

*Figura 3. Segmentazione Nielsen*



*Fonte: Elaborazione personale, PowerPoint*

- REGIONE: è una variabile che indica la regione in cui si trova l'agenzia di riferimento del cliente (tenere bene a mente che i clienti online non hanno una regione di riferimento)
- DATA INTERVISTA: data in cui è stato svolto il questionario (dal 24 marzo al 31 dicembre 2020);
- NPS(TM): voto da 0 a 10 dato dal cliente sulla raccomandabilità dell'azienda nel suo complesso;
- VERBATIM: commento positivo o negativo sul voto dato;
- MOTIVAZIONI: tutti i *drivers* menzionati da un cliente relativamente alla propria esperienza con la banca;
- FEDELTÀ: è una variabile che indica la volontà del cliente di voler proseguire o interrompere il rapporto, ha tre possibili risposte:
  1. Intenzione di proseguire il rapporto;
  2. Intenzione di chiudere il rapporto nei prossimi 12 mesi;
  3. Intenzione di valutare altre offerte di altre Banche nei prossimi 12 mesi;

Oltre a queste variabili già presenti nel dataset, per lo scopo di questo elaborato sono state ricavate ulteriori variabili, alcune di fondamentale importanza poiché cuore del tema fin qui proposto, ossia:

- Rabbia: variabile dicotomica che assegna valore 1 se nella variabile Verbatim della rispettiva osservazione il cliente esprime “rabbia”, 0 altrimenti;
- Disgusto: variabile che assume valori 0 e 1, valore 1 se il Verbatim esprime “disgusto”, 0 altrimenti;
- Gioia: valore 1 “gioia”, 0 altrimenti;
- Tristezza: valore 1 “tristezza”, 0 altrimenti;
- Paura: valore 1 “paura”, 0 altrimenti;
- Sorpresa: valore 1 “sorpresa”, 0 altrimenti;

### 2.2.1 Qualità del dato

Nel paragrafo precedente è stato presentato il dataset di 45713 osservazioni, campione estratto da una popolazione di 1800000 individui, con frazione di campionamento pari a 0,025 (45713/1800000).

Il campione fornito dalla banca è frutto di un disegno campionario che segue un processo a più fasi. Inizialmente viene selezionato un campione stratificato, usando come strati la segmentazione *personas*, di conseguenza la percentuale di ogni strato nel campione (ciascun segmento) rispecchia la percentuale di ogni strato nella popolazione.

Una volta stabiliti gli strati, in ciascun strato è stato fatto un campionamento casuale semplice per determinare i clienti a cui mandare il questionario. A questi viene inviata una mail con un link diretto. Il tasso di partecipazione medio dei clienti è di circa il 10%. Un problema è che, una volta inviato correttamente il questionario a tutti i clienti estratti, non è detto che questi rispondano ed infatti c'è un leggero sbilanciamento, dato che le persone più *smart* e *digitali* tendono a rispondere di più dei *tradizionali*.

Non conosciamo la percentuale di campionamento sulle singole *personas*, dato che viene ritenuto sensibile: è da presumere, come detto, una maggiore rappresentatività dei segmenti più giovani rispetto a quelli più anziani.

## 2.3 Analisi esplorativa univariata

Dopo aver dato uno sguardo alle variabili che compongono il *dataset* a disposizione, è necessario partire con un'analisi esplorativa, in particolar modo di quelle fondamentali per lo sviluppo di tale elaborato.

Il primo grafico mostra l'andamento delle 45.713 interviste effettuate tramite questionario online nel periodo, che come già detto, va dal 24 marzo 2020 al 31 dicembre 2020.

*Grafico 1. Distribuzione interviste*

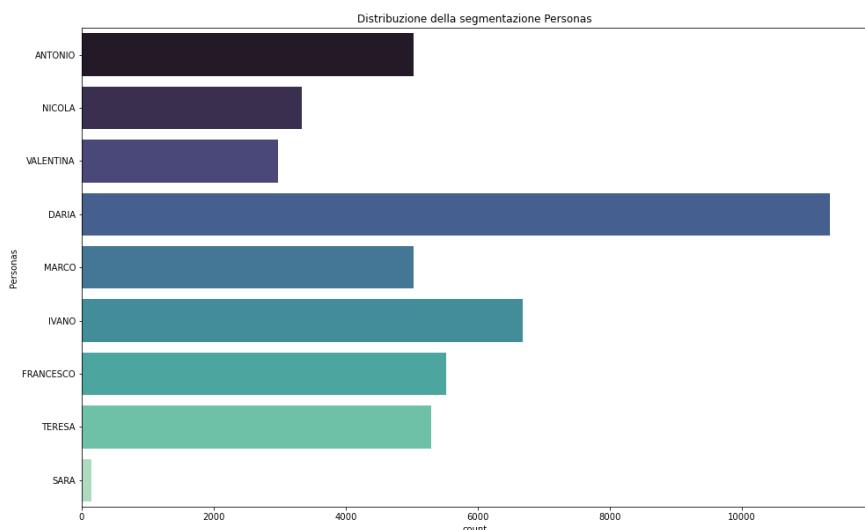


*Fonte: Elaborazione personale, Excel*

Dal grafico si può facilmente notare che nel primo periodo, escludendo marzo che comprende solo l'ultima settimana del mese, il numero delle interviste mensili tende ad essere abbastanza costante sopra i 4.000, numero che poi, con l'arrivo della stagione estiva tende pian piano a diminuire fino ai 2.660 di agosto, per poi aumentare di molto

con l'arrivo della stagione autunnale e, quindi, con la riapertura di tutte le attività lavorative. Dopo aver visto la distribuzione delle interviste, risulta interessante vedere la distribuzione della segmentazione “*Personas*”, che si rifa sugli standard della segmentazione *Eurisko*<sup>28</sup> nel mondo finanziario.

Grafico 2. Distribuzione della segmentazione Personas



Fonte: Elaborazione personale, Python

La prima cosa che salta fuori è la poca presenza del segmento “Sara”; tuttavia, ciò è abbastanza naturale e logico perché si tratta di un segmento molto giovane, dai 19 ai 25 anni di età, classe in cui i ragazzi sono ancora impegnati con il mondo universitario e più in generale dello studio.

Anche l'alta presenza di “Daria” è giustificata perché comprende la fascia di clientela con bassa scolarizzazione, medio/basso reddito, quindi tutte quelle persone che

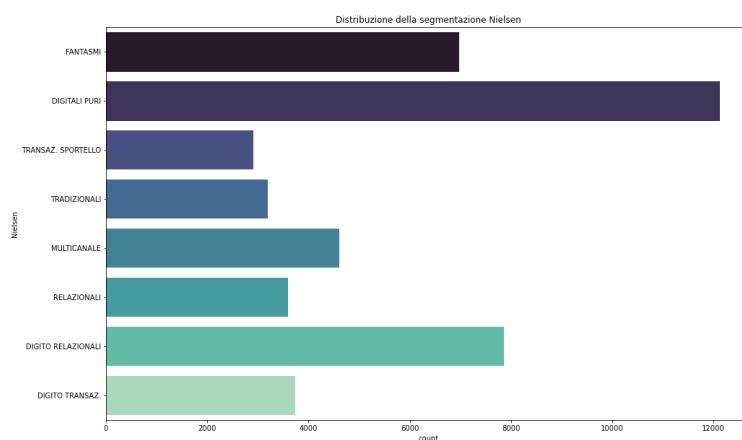
---

<sup>28</sup> Una segmentazione fatta sulla base di tre variabili: reddito, età e livello d'istruzione.

svolgono lavori umili (commessi, operai, etc).

Un'altra segmentazione di rilevante importanza è la segmentazione *Nielsen* (campione di 45.009 osservazioni), anch'essa ottenuta (in una precedente analisi) tramite la combinazione di tre variabili, rispettivamente “*Sportello*” (aver fatto, o no, operazioni di sportello nell'ultimo anno), “*Gestore*” (aver avuto o no contatti con un gestore, come magari contatti di consulenza, nell'ultimo anno) e “*Internet*” (cliente attivo in modo digitalizzato, effettuando operazioni *online*).

*Grafico 3. Distribuzione della segmentazione Nielsen*



*Fonte: Elaborazione personale, Python*

Notiamo come i “*digitali puri*” siano in prevalenza numerica rispetto agli altri segmenti, seguiti da “*digito-relazionali*”, ed oggi, in un mondo che diventa sempre più digitalizzato e all'avanguardia, una cosa del genere è abbastanza scontata; interessante è l'alta presenza di “*fantasmi*”, ossia clienti che non hanno avuto nessun tipo di contatto

con la banca negli ultimi 12 mesi.

Infine, si veda come i segmenti che non hanno effettuato nemmeno un accesso online, tra cui “*relazionali*”, “*tradizionali*” e “*transazionali-sportello*” siano in minoranza, dimostrazione del fatto che oggi la maggior parte delle operazioni si svolgono online e che, agenzia e sportelli vengono utilizzati prevalentemente dalla clientela più *old* che ha meno familiarità con la tecnologia.

Dopo aver visto le varie distribuzioni del campione in base alle diverse tipologie di segmentazione della banca, è stato necessario soffermarsi anche sulla distribuzione geografica delle osservazioni sul territorio italiano, facendo attenzione al fatto che in questo caso la numerosità del campione è scesa a 34.749, poiché sono state prese in considerazione solamente le osservazioni appartenenti ad un'agenzia e non quindi i clienti che hanno contatti con la banca online.

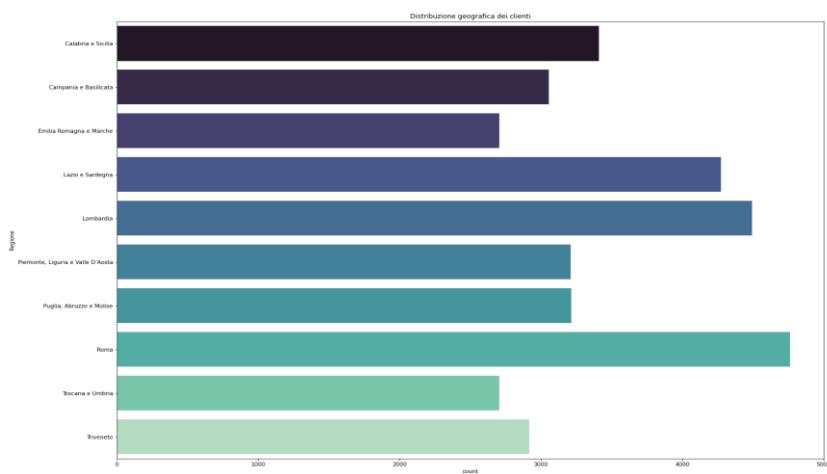
*Grafico 4. Distribuzione geografica dei clienti*



*Fonte: Elaborazione personale, Excel – Le tonalità più scure indicano le regioni dove la percentuale dei rispondenti è risultata maggiore*

Volendo accorpate alcune tra di loro, le regioni più simili e vicine, seguendo la suddivisione adottata dalla banca, si ottiene una distribuzione di frequenza di questo tipo:

*Grafico 5. Distribuzione geografica dei clienti*



*Fonte: Elaborazione personale, Python*

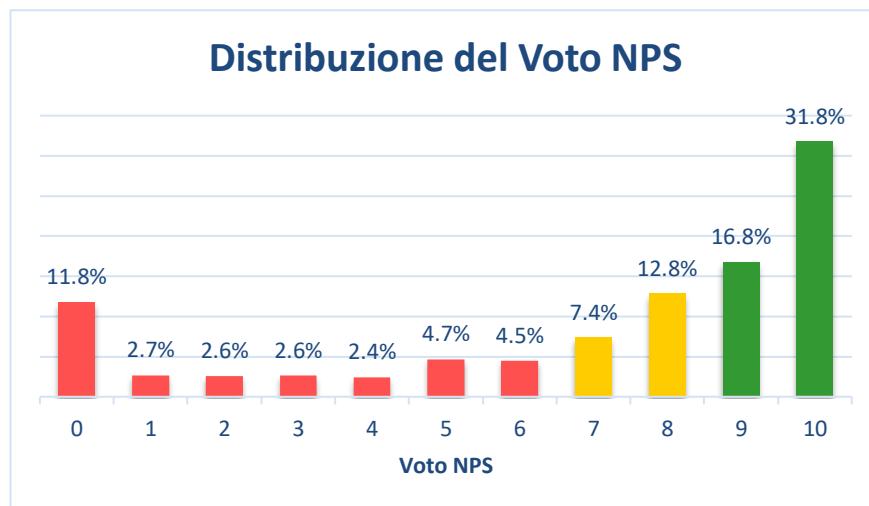
Vediamo come le frequenze assolute più alte delle osservazioni si trovano a “Roma”, in “Lombardia” e “Lazio e Sardegna”, ossia le regioni che comprendono le città italiane con più disponibilità di servizi e lavoro in generale, Milano e Roma; seguono poi “Calabria e Sicilia”, “Piemonte, Liguria e Valle d’Aosta” e tutte le altre.

### 2.3.1 Le Analisi del Net Promoter Score (NPS)

Come già ampiamente detto nel primo capitolo, l'analisi del *Net Promoter Score*, o più semplicemente *NPS*, ha preso velocemente piede negli ultimi 20 anni, a causa della sua versatilità (può essere applicata in tutti gli ambiti) e delle informazioni che permette di restituire.

Per prima cosa, è conveniente andare a visualizzare la distribuzione di frequenze relative dei voti *NPS* dati dai clienti presenti nel campione per poter iniziare a farsi un'idea generale dell'andamento della banca durante questo periodo marzo/dicembre.

Grafico 6. Distribuzione Voto NPS



Fonte: Elaborazione personale, Excel

Si può notare come la distribuzione si concentri particolarmente sui voti più alti,

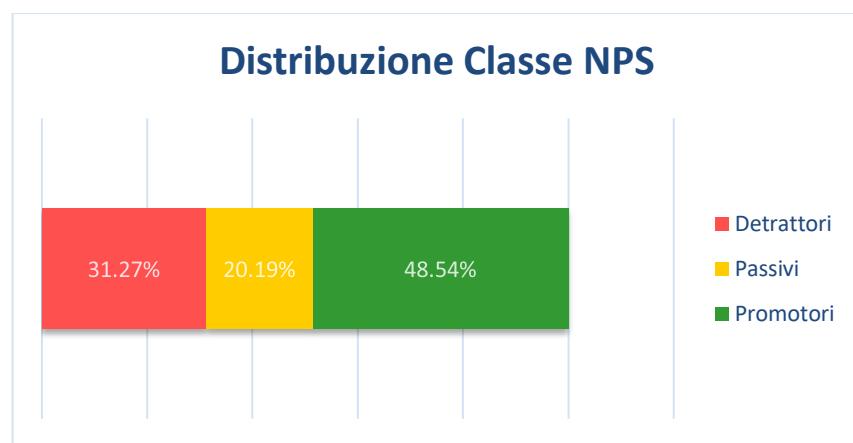
dall'otto a salire, cosa che fa particolarmente piacere poiché sinonimi di un buon stato di salute della banca: almeno in modo generale le attività sono state svolte bene.

I promotori (*Promoters*), che ricordiamo essere coloro che hanno espresso un voto dal 9 in su, sono circa il 48,6% a cui corrispondono, in valore assoluto, ben 22190 clienti.

L'altra parte della medaglia, i detrattori (*Detractors*) che hanno espresso un voto compreso tra 0 e 6, conta il 31,3% circa, a cui corrispondono 14.293 clienti.

In questo caso, bisogna fare molta attenzione a quei detrattori che hanno espresso il voto 0, perché 11,8% è abbastanza elevato e, soprattutto, coloro che esprimono i voti più bassi sono i clienti più insoddisfatti di tutti, clienti “arrabbiati” per un qualsiasi motivo e la cosa più importante è che questi clienti non solo sono pronti a lasciare la banca da un momento all'altro, ma potrebbero effettuare all'esterno una sorta di pubblicità negativa che ovviamente può solo danneggiare la banca.

Grafico 7. Distribuzione Classe NPS



Fonte: Elaborazione personale, Excel

La figura sintetizza il grafico precedente sulla distribuzione dei voti, accorpando ciascun voto nella sua *Classe NPS* di appartenenza (detrattore, passivo o promotore).

La pubblicità negativa fatta dai clienti stessi è quella che viene ascoltata di più, basti pensare alle recensioni che spesso vengono lette su siti come *Amazon*, *TripAdvisor*; spesso si tende a dare più importanza alle risposte negative che a quelle positive.

Per capire quanto è stato positivo/negativo il lavoro della banca bisogna calcolare l'*NPS* come differenza:

$$\textcolor{green}{\% \text{Promotori}} - \textcolor{red}{\% \text{Detrattori}}$$

Il risultato va poi moltiplicato \*100 per non avere una percentuale ma un numero puro.

In questo caso l'*NPS* è:

$$(48.54\% - 31.27\%)*100$$

$$= 17.27^{29}$$

---

<sup>29</sup> Qualità dell'indice: Il valore dell'*NPS* 17.27 ha l'intervallo di confidenza al 95%, [17.13; 17.41]. La formula usata per il calcolo è la seguente:  $\text{NPS} \pm \sqrt{\frac{p(1-p)+d(1-d)+2pd}{N-1}}$  dove  $p$  e  $d$  sono rispettivamente le percentuali di *promotori* e *detrattori*,  $N$  è la numerosità del campione.

Quel 17.27, come detto su, è stato moltiplicato \*100, in modo da togliere la percentuale.

La banca, quindi, ha ottenuto un punteggio assolutamente positivo, in quanto va ritenuto buono qualsiasi punteggio al di sopra dello 0, dimostrando il benessere generale della stessa nel periodo preso in analisi.

Purtroppo, non basta fermarsi qui, sarebbe fantastico poter dire che le cose vanno bene semplicemente con una banale differenza, ma sarebbe come giudicare un libro dalla copertina; infatti, bisogna andare a cercare più nelle profondità della banca per evidenziare al meglio le relazioni utili per lo scopo di questo elaborato.

La promozione/detrazione è un tema molto forte che caratterizza le persone, ma è ben risaputo che ognuno reagisce ad una determinata circostanza a proprio modo, diverso dagli altri; quindi, è fondamentale iniziare vedendo delle differenze tra la clientela della banca, clientela suddivisa in più segmentazioni.

La prima, basata sulle *Personas*, ci restituisce questi risultati:

*Tabella 1.*

Personas/Classe NPS	Detrattore	Passivo	Promotore	Totale
ANTONIO	0.34	0.22	0.44	1.00
DARIA	0.33	0.18	0.49	1.00
FRANCESCO	0.33	0.19	0.49	1.00
IVANO	0.25	0.25	0.51	1.00
MARCO	0.36	0.18	0.45	1.00
NICOLA	0.23	0.22	0.55	1.00
SARA	0.30	0.16	0.54	1.00
TERESA	0.20	0.23	0.56	1.00
VALENTINA	0.45	0.17	0.37	1.00
<b>Totale</b>	<b>0.31</b>	<b>0.20</b>	<b>0.49</b>	<b>1.00</b>

Fonte: Elaborazione personale, Excel – Sono colorati i valori superiori alla media

La tabella è molto interessante perché ci mostra la distribuzione di frequenza congiunta della segmentazione *Personas*, andando a vedere per ciascuna tipologia di persona se la distribuzione di detrattori, passivi e promotori al suo interno è uguale alla distribuzione di frequenza marginale. In parole povere si va a confrontare la distribuzione “totale” con quella di ciascuna “persona”.

Risulta ben evidente come alcune persone (*Ivano, Nicola, Sara e Teresa*) abbiano una percentuale di promotori maggiore rispetto a quella “totale”, il che potrebbe indicare che tali classi di persone siano più inclini all’essere promotori (per *Ivano, Nicola* e *Teresa* vi è anche una bassa percentuale di detrattori). Escludendo *Sara*, si tratta dei segmenti più anziani, più tradizionali, che costituiscono, da sempre una delle componenti più importanti della *Customer Base*.

Importante per la banca sono, però, i detrattori, poiché sono coloro che non portano valore, coloro che distruggono valore e che al di fuori della sfera bancaria parlano male della stessa.

Il più significativo è *Valentina*, che conta ben 45% di detrattori, percentuale davvero molto alta rispetto alle altre, meno rilevanti ma comunque importanti, come *Marco* 36%, *Antonio* 34%, *Daria* e *Francesco* 33%.

Andando a controllare la composizione della segmentazione si può notare come la classe over 60 (*Nicola, Teresa* e *Ivano*) ha tendenzialmente espresso voti più elevati, mentre la classe compresa tra 26 e 60 anni (*Valentina, Antonio, Marco, Francesco* e

*Daria*) ha avuto meno soddisfazione, optando per voti più negativi.

Un *focus* importante su *Valentina* (classe 26-45 anni, alta scolarizzazione e reddito alto) la cui classe è stata oggetto di detrazione forte (solo 37% di promotori rispetto al 49% di media).

Questa differenziazione è dovuta al fatto che in ogni classe vi sono particolari bisogni e problematiche.

La stessa distribuzione congiunta è stata fatta anche sulla segmentazione *Nielsen*:

*Tabella 2*

Nielsen/Classe NPS	Detrattore	Passivo	Promotore	Totale
DIGITALI PURI	0.31	0.23	0.46	1.00
DIGITO RELAZIONALI	0.36	0.20	0.45	1.00
DIGITO TRANSAZ.	0.29	0.22	0.49	1.00
FANTASMI	0.30	0.18	0.51	1.00
MULTICANALE	0.29	0.20	0.51	1.00
RELAZIONALI	0.33	0.18	0.49	1.00
TRADIZIONALI	0.25	0.17	0.57	1.00
TRANSAZ. SPORTELLO	0.26	0.19	0.55	1.00
<b>Totale</b>	<b>0.31</b>	<b>0.20</b>	<b>0.49</b>	<b>1.00</b>

*Fonte: Elaborazione personale, Excel – I valori colorati sono quelli maggiori della media totale*

Riguardo i detrattori, si nota come quelli più propensi risultano essere *Digito Relazionali* e *Relazionali*, ossia clienti che nell'ultimo anno non hanno avuto contatti con lo “*Sportello*” di una banca ma che, invece, hanno avuto contatti con un “*Gestore*” (per esempio un consulente); per i promotori, le percentuali più alte, in ordine decrescente, si hanno nei *Tradizionali*, *Transazionali Sportello*, *Fantasmi* e *Multicanale*, dove i primi tre hanno in comune il fatto di non essere clienti che hanno avuto contatti su “*Internet*”.

Infine, è stata vista anche la distribuzione congiunta tra *Classe NPS* e la *Regione*:

*Tabella 3*

Regione/ Classe NPS	Detrattore	Passivo	Promotore	Totale
<b>Calabria e Sicilia</b>	0.27	0.18	<b>0.55</b>	1.00
<b>Campania e Basilicata</b>	<b>0.35</b>	0.19	0.46	1.00
<b>Emilia Romagna e Marche</b>	0.29	<b>0.22</b>	0.48	1.00
<b>Lazio e Sardegna</b>	<b>0.32</b>	0.20	0.49	1.00
<b>Lombardia</b>	<b>0.32</b>	<b>0.22</b>	0.45	1.00
<b>Piemonte, Liguria e Valle D'Aosta</b>	<b>0.31</b>	<b>0.22</b>	0.47	1.00
<b>Puglia, Abruzzo e Molise</b>	0.30	0.19	<b>0.51</b>	1.00
<b>Roma</b>	<b>0.32</b>	<b>0.21</b>	0.47	1.00
<b>Toscana e Umbria</b>	0.29	<b>0.21</b>	<b>0.51</b>	1.00
<b>Triveneto</b>	0.25	<b>0.21</b>	<b>0.53</b>	1.00
<b>Totale</b>	<b>0.30</b>	<b>0.21</b>	<b>0.49</b>	<b>1.00</b>

*Fonte: Elaborazione personale, Excel*

Nella tabella risultano molto evidenti “*Calabria e Sicilia*” e “*Triveneto*” tra le regioni con tasso più alto di promotori (allo stesso tempo tasso più basso di detrattori) seguite da “*Toscana e Umbria*” e “*Puglia, Abruzzo e Molise*”, mentre, per quanto riguardo i detrattori la percentuale più significativa si ha per “*Campania e Basilicata*”.

### 2.3.2 Cos'è successo durante il periodo *Covid-19*?

Dopo aver analizzato la composizione di detrattori, passivi e promotori rispetto alle varie segmentazioni e alla posizione geografica, nasce spontaneamente una domanda:

“Durante il periodo marzo-dicembre 2020, quindi dall’arrivo del *Covid-19* in Italia fino alla Seconda Ondata, nella banca le cose sono state sempre le stesse?”

Si va ad analizzare, quindi, la percentuale delle varie *Classi NPS* per ogni mese del periodo in esame:

*Tabella 4.*

Mese/ Classe NPS	Detrattore	Passivo	Promotore	Totale
<b>Marzo</b>	0.29	0.19	0.52	1.00
<b>Aprile</b>	0.32	0.21	0.48	1.00
<b>Maggio</b>	0.33	0.20	0.47	1.00
<b>Giugno</b>	0.35	0.20	0.45	1.00
<b>Luglio</b>	0.35	0.20	0.45	1.00
<b>Agosto</b>	0.34	0.19	0.46	1.00
<b>Settembre</b>	0.31	0.21	0.48	1.00
<b>Ottobre</b>	0.30	0.20	0.49	1.00
<b>Novembre</b>	0.29	0.21	0.50	1.00
<b>Dicembre</b>	0.28	0.20	0.52	1.00
<b>Totale</b>	<b>0.31</b>	<b>0.20</b>	<b>0.49</b>	<b>1.00</b>

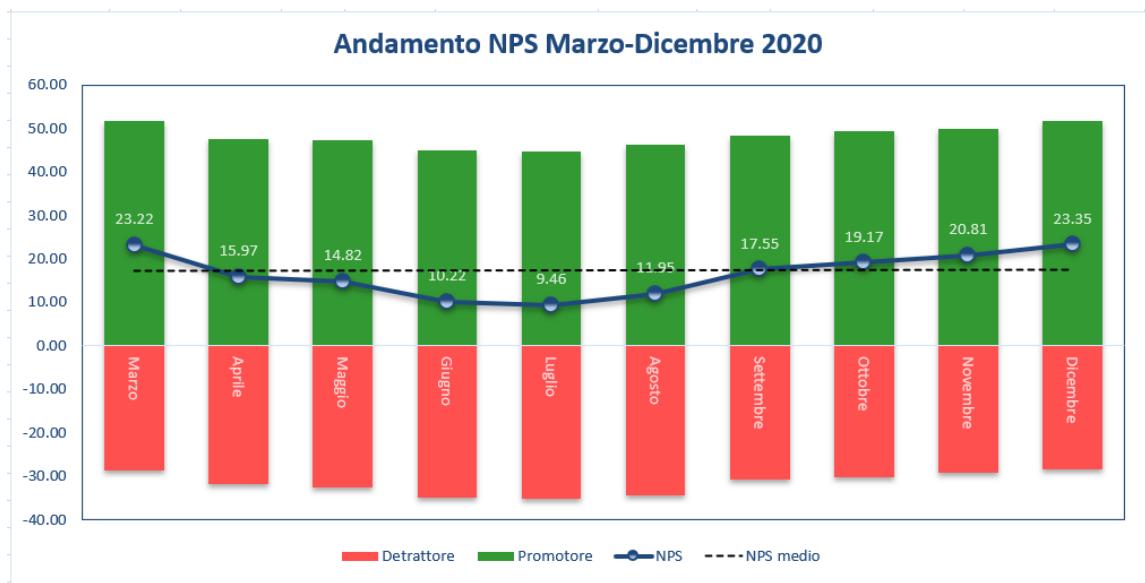
*Fonte: Elaborazione personale, Excel*

Qui si inizia ad intravedere una disparità tra i mesi; escludendo marzo (intervista solamente sugli ultimi sette giorni), notiamo nei primi mesi un aumento dei detrattori

fino alla fine della stagione estiva. Successivamente ricominciano a salire i promotori.

Per avere un quadro ancora più dettagliato della situazione, mettiamo a confronto gli *NPS* di ogni mese con l'*NPS* medio calcolato prima.

Grafico 8. Andamento NPS periodo Marzo-Dicembre 2020



Fonte: Elaborazione personale, Excel

Il grafico mostra in modo ancora più preciso la situazione che già si poteva notare nella tabella precedente.

Si parte da un valore dell'*NPS* pari a 23.22 di marzo con successivi decrementi fino ad arrivare al minimo di 9.46 di luglio, la situazione poi da agosto in poi sembrerebbe ristabilizzarsi con un forte incremento proprio tra agosto e settembre (da 11.95 a 17.55) arrivando al 23.35 di dicembre.

La situazione è diventata importante nel periodo aprile-luglio, periodo che coincide in pieno con la pandemia *Covid-19* e che nell'ambito della banca ha visto una decrescita

dei promotori a favore di una crescita dei detrattori che si traduce in un aumento della insoddisfazione e del malessere dei clienti, per questo motivo bisogna andare a vedere più nello specifico cosa ha sbagliato la banca in questi mesi difficili per tutti.

### 2.3.3 Le “colpe” della Banca

È stato visto che nel periodo aprile-luglio la banca ha avuto una decrescita in termini di *NPS* e nonostante il periodo storico, che è stato molto difficile, la banca stessa ha le sue colpe.

Per capire quali sono state le sfere maggiormente prese di mira dalla detrazione, si va ad analizzare le “*motivazioni*”, sempre per *Classe NPS*, facendo però un confronto tra marzo (mese di inizio in cui ancora non si erano verificati gli effetti della pandemia) e i restanti mesi in termini di scostamento, i cui risultati sono:

*Tabella 5.*

	Detratore	Passivo	Promotore
<b>Accessibilità/Reperibilità</b>	-0.01	-0.01	0.02
<b>Accesso al credito</b>	-0.11	-0.02	0.13
<b>Affidabilità della banca</b>	0.07	-0.01	-0.06
<b>Affidabilità/sicurezza app</b>	0.23	0.18	-0.41
<b>Affidabilità/sicurezza sito</b>	-0.03	-0.12	0.16
<b>Ambiente filiale</b>	0.07	-0.24	0.18
<b>Assenza sportello</b>	0.12	-0.14	0.02
<b>ATM</b>	0.20	-0.15	-0.06
<b>Bancomat/Carta di credito</b>	0.13	-0.09	-0.04
<b>Burocrazia</b>	0.02	0.09	-0.11
<b>Capillarità</b>	0.11	-0.13	0.02
<b>Chiarezza informazioni</b>	-0.10	-0.02	0.11
<b>Competenza/Professionalità</b>	0.01	-0.01	0.00
<b>Costi</b>	0.11	-0.02	-0.09
<b>Costi in filiale</b>	0.01	0.26	-0.27
<b>Cura del cliente</b>	0.03	0.03	-0.06
<b>Disponibilità del sito</b>	0.07	-0.22	0.15
<b>Disponibilità dell'app</b>	0.18	0.09	-0.26
<b>Disponibilità/Cortesia</b>	0.01	-0.01	-0.01
<b>Gamma dei prodotti</b>	-0.04	-0.02	0.06
<b>Immagine della banca</b>	-0.03	0.08	-0.05
<b>Innovatività</b>	0.20	0.22	-0.43
<b>Modifica condizioni</b>	0.03	-0.05	0.02
<b>Orari filiale</b>	-0.08	0.04	0.04
<b>Problem solving/Proattività</b>	-0.05	-0.03	0.08
<b>Qualità del servizio</b>	-0.01	0.05	-0.04
<b>Relazione con DAG</b>	0.17	-0.15	-0.02
<b>Rendimento prodotti</b>	-0.15	0.20	-0.05
<b>Serietà/correctezza</b>	-0.03	-0.01	0.04
<b>Servizio assistenza chat</b>	-0.02	-0.07	0.09
<b>Servizio assistenza telefonica</b>	-0.02	0.08	-0.06
<b>Tempi di attesa in agenzia</b>	0.05	-0.06	0.01
<b>Tempi di risoluzione</b>	-0.15	0.01	0.14
<b>Trasparenza, chiarezza</b>	0.06	-0.07	0.01
<b>Usability (generico)</b>	0.14	0.10	-0.24
<b>Usability app</b>	0.02	-0.01	0.00
<b>Usability sito</b>	0.05	-0.02	-0.03
<b>Totale complessivo</b>	0.03	0.00	-0.03

Fonte: Elaborazione personale, Excel

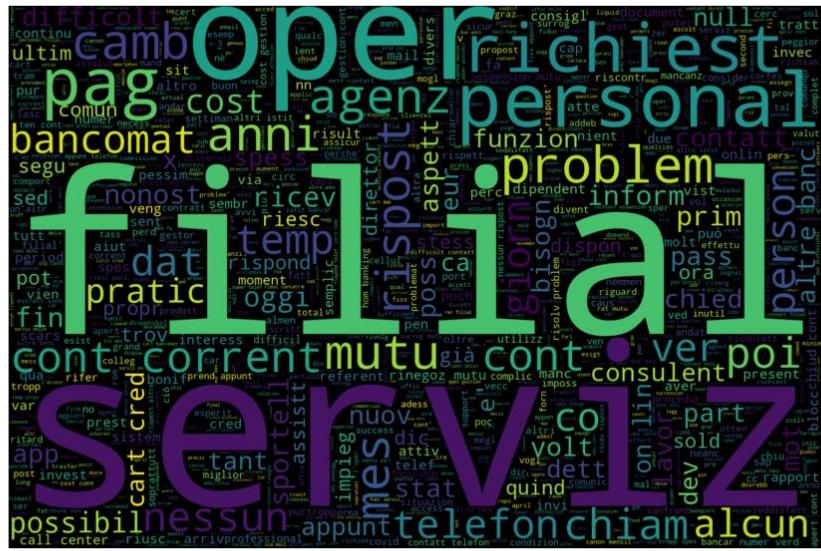
I valori della tabella esprimono l'aumento percentuale di ogni classe *NPS* tra marzo e i mesi successivi, per cercare di capire che effetti ha avuto il *Covid-19*. Come si può ben intuire le tematiche di detrazione sono molte, le quali possono essere suddivise in macro-tematiche; quelle che saltano più all'occhio sono sicuramente “affidabilità/sicurezza app”, “ATM”, “disponibilità app”, “relazione con DAG”, “usability”, “assenza sportelli” e così via.

Alcuni dei fattori critici, tra quelli che hanno avuto gli incrementi più importanti di detrazione, sono collegati all'app, magari lente, non molto *user-friendly*; bisogna però sottolineare il fatto che durante i mesi di pandemia le reti *internet* sono notevolmente calate (in termini di velocità) a causa dell'enorme massa di utenza collegata contemporaneamente, poiché tutte le persone erano costrette a restare chiuse in casa, ed infatti notiamo che anche la “*disponibilità del sito*” ha avuto un aumento, in termini di percentuale di detrattori, da marzo in poi.

Importante è anche la tematica dell'assenza di sportelli, poiché alcuni sono stati chiusi proprio a causa del *Covid*, che come ben sappiamo ha dato via a numerose limitazioni, così come “*relazione con DAG*”, dato che i clienti cercavano di mettersi in contatto e di prendere un appuntamento con consulenti e direttori, ovviamente con scarsi risultati perché il periodo non aiutava di certo.

Tutte queste problematiche vengono confermate anche dallo studio dei *Verbatim* dei soli detrattori nel periodo marzo-dicembre; infatti, andando a vedere le parole più usate risultano:

*Figura 4. WordCloud detrattori marzo-dicembre*



*Fonte: Elaborazione personale, Python*

La *WordCloud*<sup>30</sup> conferma quanto visto sopra con la tabella, mostrando le “parole chiave” dei commenti dei detrattori, tra cui spuntano proprio “*serviz*”<sup>31</sup> (servizio/i), “*filial*” (filiale/i), conferma che i servizi non sono stati dei migliori, così come le problematiche riguardo le filiali.

Bisogna tenere a mente però, che il periodo storico di riferimento è stato molto particolare, è vero che le problematiche sono state tante e che sicuramente la banca ha le

<sup>30</sup> WordCloud è una rappresentazione visiva di etichette o parole-chiave usate, in questo caso, nei questionari dei clienti.

<sup>31</sup> Attraverso un processo chiamato *stemming* le parole perdono la loro desinenza, questo per non considerare, per esempio, “servizio” e “servizi” come due parole differenti.

sue colpe, i suoi limiti, ma è anche vero che durante questo periodo il paese si è trovato in un *lockdown*, una chiusura nazionale di tutte le attività e che ha costretto le persone a restare chiuse in casa e questo ha implicato un sovraccarico delle linee internet che ha notevolmente rallentato il caricamento di “siti” e “app”.

Stessa cosa per quanto riguarda magari i “*tempi di attesa in un’agenzia*” o “*relazione con DAG*”; potevano entrare solo un numero limitato di persone alla volta con annesso distanziamento e ciò favoriva il crearsi di code chilometriche. Era diventato tutto più difficile, anche solamente prendere un appuntamento con un consulente perché la banca, ma in generale il mondo intero recente non aveva mai affrontato una situazione del genere, di questa portata, e di conseguenza tutti i settori si sono trovati presi alla sprovvista.

Le tematiche toccate dai detrattori sono svariate, ma chi sono realmente questi detrattori? Quale segmento ha reagito peggio? E meglio?

Ciò che ne segue è un’analisi della relazione tra le motivazioni, raggruppate in alcune macro-tematiche (per una più facile lettura e comprensione), e i vari segmenti di clientela, per cercare di capire l’atteggiamento e le problematiche più diffuse di ciascuno.

Per una più semplice interpretazione è stato utilizzato il voto medio delle motivazioni condizionato ai vari segmenti:

*Tabella 6.*

	ANTONIO	DARIA	FRANCESCO	IVANO	MARCO	NICOLA	SARA	TERESA	VALENTINA	Totale
<i>Accessibilità/Reperibilità</i>	3.3	3.0	3.1	4.3	3.0	4.0	1.6	4.1	2.6	3.4
<i>Accesso al credito</i>	3.5	3.1	4.0	5.5	3.3	2.7		4.0	3.8	3.6
<i>Affidabilità della banca</i>	8.5	8.8	8.6	8.6	8.6	8.6	8.7	8.7	8.5	8.6
<i>Affidabilità/sicurezza sito/app</i>	7.8	7.2	7.7	7.3	7.2	8.7	8.5	8.4	6.6	7.5
<i>Ambiente filiale</i>	4.7	4.2	5.5	5.0	5.0	5.0	5.0	4.9	3.9	4.7
<i>Assenza sportello</i>	4.0	3.6	4.2	5.1	3.7	5.3		4.9	3.6	4.4
<i>ATM</i>	4.2	4.3	5.0	5.1	5.1	5.1		5.1	4.6	4.8
<i>Bancomat/Carta di credito</i>	2.8	3.5	3.7	4.5	3.6	4.7	5.0	4.8	2.9	3.6
<i>Burocrazia</i>	3.7	3.6	2.9	4.3	3.3	4.3		5.0	2.9	3.7
<i>Capillarità</i>	6.1	5.9	6.4	6.9	5.7	6.3		6.2	4.5	6.1
<i>Chiarezza informazioni</i>	6.0	6.1	6.6	6.8	5.7	6.1	8.5	7.3	5.1	6.3
<i>Comodità</i>	9.0	9.2	8.8	8.0	9.0	9.1	10.0	8.7	8.9	8.9
<i>Competenza/Professionalità</i>	8.2	8.1	8.5	8.8	8.1	8.8	9.1	9.0	7.3	8.4
<i>Costi</i>	5.2	5.2	5.1	5.6	5.1	5.2	6.7	5.4	4.6	5.1
<i>Cura del cliente</i>	4.9	4.1	4.4	6.0	4.2	6.2	3.4	6.5	3.8	4.9
<i>Disponibilità sito/app</i>	6.8	7.1	6.1	6.0	6.2	5.6	9.3	6.8	5.7	6.4
<i>Disponibilità/Cortesia</i>	8.0	8.0	8.3	8.6	7.8	8.6	8.9	8.8	7.3	8.3
<i>Gamma dei prodotti</i>	7.8	7.8	7.4	6.7	8.3	7.3	9.0	6.4	5.9	7.4
<i>Immagine della banca</i>	7.8	8.7	8.4	8.3	8.4	8.3	9.6	8.7	8.2	8.4
<i>Innovatività</i>	7.7	8.2	7.8	7.7	7.7	9.5		8.3	6.5	7.9
<i>Modifica condizioni</i>	3.6	3.1	3.5	3.6	4.2	3.6	0.7	4.3	3.7	3.6
<i>Problem solving/Proattività</i>	5.1	5.8	5.2	6.8	4.7	6.8	3.0	6.9	4.5	5.7
<i>Promozioni</i>	5.0	2.4	8.3	4.2	5.5	7.8	8.0	5.2	6.1	4.9
<i>Qualità del servizio</i>	7.8	8.0	7.9	8.2	7.7	8.3	7.8	8.3	7.6	8.0
<i>Relazione con DAG</i>	4.5	4.9	4.4	4.1	4.9	2.8	8.0	6.0	2.5	4.5
<i>Rendimento prodotti</i>	3.8	3.8	5.5	4.6	2.6	4.0		4.4	4.6	4.3
<i>Serietà/correctezza</i>	8.3	8.0	8.4	8.6	7.8	8.9	8.0	8.7	7.4	8.3
<i>Servizio assistenza chat</i>	5.2	5.5	7.0	5.4	6.7	6.0	10.0	9.0	5.0	6.0
<i>Servizio assistenza telefonica</i>	4.7	4.9	4.7	5.1	4.9	5.8	5.0	5.0	4.1	4.8
<i>Tempistiche</i>	3.9	4.2	4.7	5.7	4.0	5.0	3.0	5.9	4.1	4.6
<i>Trasparenza, chiarezza</i>	6.2	6.1	6.7	6.9	6.2	6.9	9.0	7.2	4.0	6.3
<i>Usability</i>	6.9	7.6	7.8	6.7	6.9	7.5	6.7	7.0	6.8	7.1
<b>Total complessivo</b>	<b>6.4</b>	<b>6.5</b>	<b>6.7</b>	<b>7.3</b>	<b>6.3</b>	<b>7.4</b>	<b>6.9</b>	<b>7.6</b>	<b>5.6</b>	<b>6.7</b>

*Fonte: Elaborazione personale, Excel*

La tabella mostra, per ciascuna *personas*, il voto medio per ogni motivazione. I colori sono stati assegnati tramite una scala colore che assegna un verde acceso al punteggio più alto (10, voto massimo) e un rosso acceso al punteggio più basso (0, voto minimo).

Dalla tabella si nota come alcune motivazioni abbiano decisamente dei voti più bassi come “accessibilità/reperibilità”, “accesso al credito”, “burocrazia”, “bancomat/carta

*di credito*”, “*modifica condizioni*”, “*servizi di assistenza telefonica*”, “*tempistiche*” e tanto altro. Fattori critici che erano già stati evidenziati in precedenza. Alcune sono distribuite in modo molto omogeneo, nel senso che sono state fonte di disprezzo per un po' tutti i segmenti tranne che per *Ivano* e *Teresa*, che sembrerebbe abbiano dato delle votazioni leggermente superiori.

Importante è la situazione di *Valentina*, un segmento quasi per natura detrattore; infatti, si può notare come la media dei voti dati è la più bassa, 5.6, molto inferiore rispetto agli altri segmenti.

Inoltre, *Valentina* sembrerebbe comportarsi in modo simile a *Marco*, e in parte anche *Francesco* e *Antonio* (classe 25-60 anni), perché la loro media voti è inferiore a quella degli altri segmenti in tematiche che riguardano specialmente il “*personale*”, come “*cura del cliente*”, “*serietà/correttezza*”, “*disponibilità/cortesia*”. Queste ultime due in particolare sono tematiche con media voto elevata ovunque, sinonimo di buon lavoro qui, però si può notare come questi segmenti avrebbero voluto qualcosa in più, qualche attenzione in più.

Per quanto riguarda la fascia over 60, si può vedere come *Teresa*, *Ivano* e *Nicola* sembrerebbero i più soddisfatti; infatti, le loro medie voto sono le più elevate, questo perché si tratta di clienti la cui maggior parte è pensionata, non hanno quindi le stesse esigenze di una persona che invece è nel pieno della propria carriera lavorativa (es. *Valentina*).

Un *Nicola* che va in agenzia e la trova chiusa reagisce in maniera diversa da una *Valentina*, poiché magari quest'ultima è impegnata, stressata con il lavoro, non ha altri momenti liberi, mentre un *Nicola*, probabilmente pensionato, ha decisamente molto più tempo per poter ritornare, dando meno peso al problema.

Tra i valori più bassi troviamo “*promozioni*” per *Daria*, questo perché si tratta di un segmento che spesso svolge lavori umili (operaio, commessi, etc), la classica persona che vorrebbe sempre qualche monetina in più, per questo alla ricerca di qualche promozione.

Successivamente si è passati a vedere i voti medi anche per quanto riguarda la segmentazione *Nielsen*:

*Tabella 7.*

	DIGITALI PURI	DIGITO RELAZIONALI	DIGITO TRANSAZ.	FANTASMI	MULTICANALE	RELAZIONALI	TRADIZIONALI	TRANSAZ. SPORTELLO	Totale
Accessibilità/Reperibilità	3.6	3.3	3.6	3.0	3.4	3.0	3.8	3.5	3.4
Accesso al credito	3.4	3.6	4.6	3.1	4.2	3.5	2.5	3.2	3.6
Affidabilità della banca	8.8	8.5	8.8	8.7	8.6	8.4	8.5	8.5	8.6
Affidabilità/sicurezza sito/app	7.2	7.7	8.1	7.7	7.8	7.2	9.5	6.7	7.5
Ambiente filiale	5.0	4.2	4.8	4.5	5.5	4.2	4.9	4.7	4.8
Assenza sportello	4.7	3.9	4.6	4.4	4.3	3.9	4.4	4.7	4.4
ATM	5.0	4.9	5.5	4.8	5.0	4.2	3.6	4.5	4.8
Bancomat/Carta di credito	3.8	3.6	4.7	3.0	3.8	3.1	3.6	2.8	3.6
Burocrazia	3.8	3.3	4.1	3.9	4.3	3.1	3.7	3.4	3.7
Capillarità	6.4	6.2	6.3	6.0	6.1	5.0	6.1	6.0	6.1
Chiarezza informazioni	6.3	5.7	6.3	6.0	6.7	6.7	6.7	7.1	6.3
Comodità (generico)	9.0	8.9	9.1	9.0	9.0	8.7	9.0	7.2	8.9
Competenza/Professionalità	8.1	8.3	8.5	8.2	8.8	8.5	8.8	8.7	8.4
Costi	5.5	5.1	5.5	4.7	5.0	4.5	4.5	4.7	5.2
Cura del cliente	4.8	4.5	4.9	4.7	5.4	4.7	5.4	5.3	4.9
Disponibilità sito/app	6.7	5.9	6.0	6.7	6.3	5.2	6.2	7.4	6.4
Disponibilità/Cortesia	8.0	8.0	8.3	8.1	8.7	8.4	8.7	8.3	8.3
Gamma dei prodotti	7.1	7.2	7.5	7.5	7.6	8.4	8.0	7.0	7.4
Immagine della banca	8.5	8.4	8.4	8.5	8.4	8.0	8.6	8.7	8.4
Innovatività	7.7	7.5	8.6	7.9	8.5	8.1	10.0	8.3	7.9
Modifica condizioni	3.8	4.0	2.9	3.3	4.0	2.9	1.4	1.8	3.6
Problem solving/Proattività	6.2	5.1	5.7	4.7	5.5	5.5	7.0	6.5	5.7
Promozioni	5.4	3.7	7.0	4.7	4.4	10.0	5.0		4.9
Qualità del servizio	8.1	7.8	8.2	7.9	8.1	7.8	7.6	8.1	8.0
Relazione con DAG	3.9	4.5	5.3	3.7	5.8	4.1	5.3	4.2	4.5
Rendimento prodotti	5.0	4.1	2.6	3.9	4.3	3.7	5.0	4.7	4.3
Serieta/correctezza	8.2	8.2	8.5	8.4	8.5	8.0	8.6	8.6	8.3
Servizio assistenza chat	6.3	4.6	5.6	8.3	7.3		10.0		6.0
Servizio assistenza telefonica	5.4	4.2	5.2	4.8	4.5	3.9	4.5	5.3	4.8
Tempistiche	4.9	4.0	4.7	4.2	4.7	4.5	5.2	5.1	4.6
Trasparenza, chiarezza	6.3	5.9	7.0	5.5	7.3	6.5	7.0	6.4	6.3
Usability	7.3	6.9	7.2	7.4	6.9	6.8	6.2	7.1	7.2
<b>Totale complessivo</b>	<b>6.7</b>	<b>6.3</b>	<b>6.9</b>	<b>6.6</b>	<b>7.0</b>	<b>6.6</b>	<b>7.3</b>	<b>7.1</b>	<b>6.7</b>

Fonte: Elaborazione personale, Excel

Dalla tabella possiamo vedere come la problematica “*ATM*”, ovviamente, sia più forte

per quei segmenti che non fanno uso del canale “*internet*”. Al contrario, problematiche

come “*problem solving/proattività*” sembrerebbero più presenti per i segmenti online.

Un'altra evidenza abbastanza ovvia è il voto più basso riguardo il fattore critico

“*comodità*” dato dai “*transazionisti sportello*”, poiché non è un canale che permette di

avere tutte le comodità possibili, richiede sempre il dover recarsi fisicamente in

un'agenzia e considerando il periodo poteva risultare abbastanza scomodo per molte

persone.

Il passo successivo, fondamentale, è stato vedere se vi fossero delle differenze a livello

territoriale, se quindi alcune problematiche fossero specifiche per un determinato

territorio o distribuite in modo omogeneo.

*Tabella 8.*

	Calabria e Sicilia	Campania e Basilicata	Emilia Romagna e Marche	Lazio e Sardegna	Lombardia	Piemonte, Liguria e Valle D'Aosta	Puglia, Abruzzo e Molise	Roma	Toscana e Umbria	Triveneto	Totale
<i>Accessibilità/Reperibilità</i>	3.3	3.2	3.6	2.9	3.4	3.5	3.7	3.3	3.7	3.8	3.4
<i>Accesso al credito</i>	2.6	2.7	3.7	3.6	4.1	3.8	2.3	3.9	3.8	4.8	3.6
<i>Affidabilità della banca</i>	8.8	8.6	8.6	8.5	8.6	8.1	8.8	8.5	8.7	8.7	8.6
<i>Affidabilità/sicurezza sito/app</i>	8.3	7.0	5.9	7.3	6.3	6.7	8.1	8.4	6.2	6.9	7.3
<i>Ambiente filiale</i>	5.1	4.0	4.3	4.1	5.0	5.0	5.3	5.6	4.6	4.9	4.8
<i>Assenza sportello</i>	4.6	4.0	4.8	4.0	4.3	5.0	3.7	4.7	5.2	4.7	4.4
<i>ATM</i>	4.2	5.2	5.5	3.6	5.1	5.6	5.2	5.1	5.0	6.4	4.9
<i>Bancamat/Carta di credito</i>	3.1	2.9	2.8	2.9	3.1	3.3	4.5	3.3	4.6	5.0	3.4
<i>Burocrazia</i>	3.0	3.3	3.8	4.0	4.0	3.3	4.3	3.5	4.0	4.2	3.8
<i>Capillarità</i>	5.8	6.1	6.0	5.8	6.1	6.4	6.8	6.9	6.1	5.6	6.2
<i>Chiarezza informazioni</i>	6.7	6.4	6.0	6.4	5.8	6.2	6.4	6.7	6.9	7.0	6.4
<i>Comodità (generico)</i>	9.6	8.9	8.6	8.5	8.6	8.9	8.8	9.4	8.3	8.1	8.8
<i>Competenza/Professionalità</i>	8.7	8.0	8.4	8.5	8.2	8.2	8.3	8.3	8.9	8.8	8.4
<i>Costi</i>	4.8	4.5	4.8	4.4	4.5	5.1	4.7	4.7	5.2	5.2	4.7
<i>Cura del cliente</i>	4.7	4.2	5.1	4.9	4.9	5.4	5.0	4.9	5.5	6.2	5.0
<i>Disponibilità sito/app</i>	5.9	6.9	5.9	6.1	6.5	5.9	6.7	6.2	6.3	7.2	6.3
<i>Disponibilità/Cortesia</i>	8.6	7.9	8.3	8.3	8.1	8.3	8.3	8.2	8.5	8.6	8.3
<i>Gamma dei prodotti</i>	7.2	6.2	7.7	6.8	7.5	6.4	6.8	7.5	8.1	7.5	7.2
<i>Immagine della banca</i>	8.8	8.8	8.4	8.3	8.3	8.5	8.6	8.5	8.5	8.7	8.5
<i>Innovatività</i>	7.7	8.2	9.3	8.9	6.7	6.3	7.7	6.8	8.6	8.4	7.7
<i>Modifica condizioni</i>	3.2	3.4	3.7	3.8	3.7	3.5	3.8	2.5	3.4	3.5	3.4
<i>Problem solving/Proattività</i>	6.0	5.1	6.2	5.3	5.4	6.0	6.8	5.8	5.6	6.6	5.8
<i>Promozioni</i>	4.0	4.0	3.0	4.2	2.7	8.3	4.5	7.0	9.5	6.0	4.9
<i>Qualità del servizio</i>	8.7	7.7	7.9	7.7	7.6	7.6	7.9	8.1	7.7	8.1	7.9
<i>Relazioni con DAG</i>	4.3	4.2	4.8	3.1	4.0	3.2	5.5	3.9	5.9	5.5	4.4
<i>Rendimento prodotti</i>	2.7	4.4	4.3	3.4	3.9	4.2	3.9	3.8	4.6	5.3	4.1
<i>Serietà/correttezza</i>	8.6	8.2	8.4	8.2	8.4	8.0	8.6	8.2	8.6	8.3	8.3
<i>Servizio assistenza chat</i>				5.5	6.0	3.5	9.0	8.0	5.0		6.0
<i>Servizio assistenza telefonica</i>	4.8	5.4	4.8	4.6	3.4	4.0	3.7	4.9	4.0	5.0	4.5
<i>Tempistiche</i>	4.7	4.5	5.2	4.7	4.4	4.8	4.5	4.7	4.3	5.5	4.7
<i>Trasparenza, chiarezza</i>	7.2	6.5	5.6	5.9	5.3	5.4	6.3	6.4	5.9	7.9	6.2
<i>Usability</i>	7.3	7.0	7.4	6.9	6.4	6.4	7.7	6.7	6.9	6.8	6.9
<i>Totale complessivo</i>	7.1	6.4	6.8	6.5	6.4	6.6	7.0	6.6	7.0	7.3	6.7

*Fonte: Elaborazione personale, Excel*

Andando a vedere territorialmente, si nota che alcune tematiche sono più problematiche in alcune regioni, piuttosto che in altre.

Partendo dall'*affidabilità/sicurezza* del sito e dell'*app*, sembrerebbe che da *Toscana* e *Marche* in su vi siano maggiori problemi, dato che si ha un media voto nettamente più bassa rispetto alle regioni meridionali. Per quanto riguarda gli “*ATM*” vediamo che la media voto più bassa si ha nella capitale, segno di una problematica più presente in questa grande città. “*disponibilità sito/app*” tende ad avere valori più bassi in alcune regioni, anche se non sembra esserci un collegamento territoriale, probabilmente è solo una problematica derivante dalla copertura delle linee internet. Importante è la “*relazione con DAG*” che in alcune regioni come “*Lazio e Sardegna*”, “*Piemonte, Liguria e Valle D'Aosta*” e “*Roma*” raggiunge livelli molto bassi, quindi, sembra che qui sia stato più difficile avere dei contatti con consulenti e direttori. Da tenere in considerazione anche “*servizio di assistenza telefonica*” con media voto più bassa in *Lombardia* e “*trasparenza e chiarezza*” che passa da regioni con media voto abbastanza bassa, a regioni con media voto abbastanza alta.

Tutto sommato, tra le problematiche che abbiamo visto essere più diffuse e che sono nate da marzo in poi, sembrerebbe esserci una certa omogeneità sul territorio italiano, a conferma del fatto che la banca si è trovata spiazzata un po' ovunque.

## 2.4 Come hanno vissuto il periodo i Clienti?

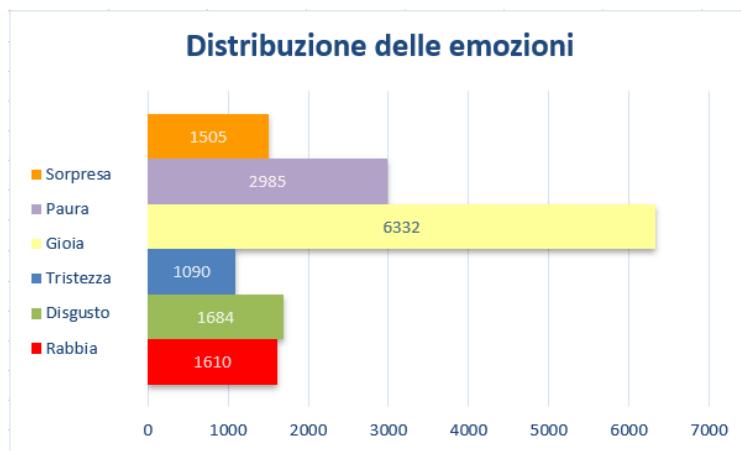
Il 2020 è stato un anno molto particolare, un anno che ha messo a dura prova il mondo intero passando dalle attività economiche (e non) fino ad arrivare alle persone stesse, costrette a dover rinunciare alla routine e adattarsi a ciò che stava succedendo.

Per le condizioni che si sono create è stato un anno unico nel suo genere e ciò ha permesso anche di poter analizzare la clientela sotto un altro punto di vista, che fino a questo punto non era mai stato preso in considerazione, quello emozionale, importantissimo proprio perché le emozioni umane, gli stati d'animo, sono stati influenzati dal mondo esterno e di conseguenza ciò può avere influito anche sulla banca.

Dai *Verbatim* contenuti nel *dataset*, grazie alla tecnica dell'*Appraisal* citata e ampiamente descritta nel capitolo precedente, è stato possibile estrarre i “*mood*” (*rabbia, tristezza, paura, gioia, disgusto e sorpresa*).

Dalle 45713 osservazioni e dai rispettivi *Verbatim* è stato possibile associare 15206 emozioni, così distribuite:

Grafico 9. Distribuzione delle emozioni



Fonte: Elaborazione personale, Excel

Il grafico<sup>32</sup> precedente mostra la distribuzione delle 15206 osservazioni a cui è stato possibile associare un'emozione, la prevalenza si ha con la “gioia” (42%), seguita da “paura” (20%), “disgusto” (11%), “rabbia” (10%), “sorpresa” (10%), “tristezza” (7%).

Come fatto anche prima, considerando che nel periodo marzo-dicembre sono successe tante cose, si va ad analizzare mese per mese questa distribuzione per vedere in che modo varia.

---

<sup>32</sup> La scelta dei colori per le emozioni non è casuale, sono stati associati gli stessi colori usati nel film *Inside Out*, con l'aggiunta della sorpresa, dove ho cercato personalmente di associare un colore adeguato.

Tabella 9.

Mesi / Mood	Rabbia	Disgusto	Tristezza	Gioia	Paura	Sorpresa	Totale
<b>Marzo</b>	0.08	0.09	0.05	0.49	0.19	0.10	1.00
<b>Aprile</b>	0.10	0.12	0.07	0.40	0.21	0.10	1.00
<b>Maggio</b>	0.11	0.10	0.08	0.40	0.22	0.10	1.00
<b>Giugno</b>	0.12	0.12	0.07	0.38	0.21	0.10	1.00
<b>Luglio</b>	0.10	0.12	0.08	0.38	0.21	0.10	1.00
<b>Agosto</b>	0.10	0.11	0.09	0.39	0.20	0.11	1.00
<b>Settembre</b>	0.10	0.10	0.08	0.44	0.17	0.11	1.00
<b>Ottobre</b>	0.09	0.10	0.07	0.44	0.19	0.11	1.00
<b>Novembre</b>	0.11	0.11	0.07	0.43	0.18	0.10	1.00
<b>Dicembre</b>	0.12	0.12	0.06	0.43	0.19	0.09	1.00
<b>Totale</b>	0.11	0.11	0.07	0.42	0.20	0.10	1.00

Fonte: Elaborazione personale, Excel – I valori colorati sono superiori alla media annuale

Questa tabella è una delle più importanti, se non la più importante, perché mostra come le varie emozioni, i vari “mood” si sono distribuiti durante l’anno, in relazione al loro peso medio annuale (ultima riga della tabella). Si nota come la “paura” sembrerebbe aumentare, rispetto alla media, ad aprile fino ad agosto, periodo che coincide proprio con il *lockdown* generale fino alle prime riaperture; lo stesso ragionamento può essere fatto per la “tristezza” che, a differenza della “paura”, si protrae anche nella stagione estiva, segno di una leggera depressione.

L’opposto si ha con la “gioia” che, durante i mesi citati sopra, diminuisce molto per poi risalire nuovamente a settembre, anche se comunque non ai livelli iniziali.

Molto interessanti sono le situazioni di “rabbia” e “disgusto” che seguono grosso modo lo stesso *trend*, con aumenti significativi tra maggio e giugno, mesi in cui ogni volta la fine del *lockdown* veniva rimandata e soprattutto novembre e dicembre, mesi in cui è stata affrontata la 2<sup>a</sup> ondata con altre stringenti restrizioni (e soprattutto restrizioni per

quanto riguarda il periodo natalizio).

Da qui l'intenzione di monitorare le emozioni mensilmente, in particolare confrontando i periodi più significati:

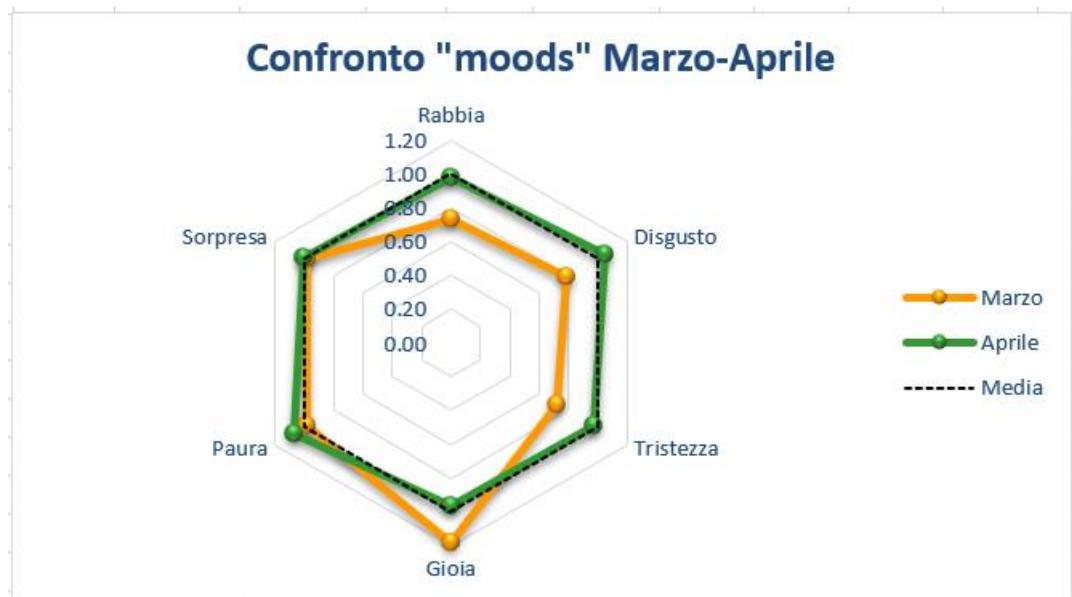
- marzo-aprile: per valutare l'impatto che ha avuto inizialmente il *Covid*;
- maggio-giugno: per vedere gli effetti della fine del *lockdown*;
- marzo-settembre: settembre è il mese della “rinascita”, dove un po' tutte le attività sono riprese, risulta interessante verificare se si è ritornati davvero ai livelli iniziali;
- settembre-dicembre: confronto tra il mese della “rinascita” e uno dei mesi più duri che ha visto altre restrizioni, in particolare per Natale, a causa della 2<sup>a</sup> ondata.

Per fare ciò è stata creata una misura normalizzata rispetto alla media, quindi dividendo ogni valore di ciascuna emozione per ogni mese per la relativa media annua (dividendo ogni riga della tabella precedente con l'ultima riga).

Bisogna precisare poi che tutti i “*moods*” non sono da attribuire solamente alle condizioni esterne, ossia a ciò che succedeva fuori dalla banca, perché già le problematiche e i punti di forza della banca stessa suscitavano un *mix* di emozioni nella clientela; ovviamente le condizioni esterne (il *Covid-19*) hanno amplificato alcuni

effetti, o meglio, hanno favorito (inevitabilmente) l'incremento di alcuni “moods”.

Grafico 10. Confronto "moods" Marzo-Aprile



Fonte: Elaborazione personale, Excel

Dal confronto tra marzo e aprile risulterebbe abbastanza evidente l'andamento di alcune emozioni, in particolare notiamo come “rabbia”, “disgusto” e “tristezza” aumentano (in modo meno marcato anche la “paura” che mese dopo mese andrà sempre aumentando); ciò è abbastanza intuitivo perché l'inizio della pandemia ha creato un *mix* di emozioni nelle persone, soprattutto negative, questo a causa delle tante restrizioni, del virus che non aveva una cura e tante altre problematiche. La “rabbia” delle persone per la situazione difficile che si stava creando, andando a sommarsi alle tante problematiche che già una famiglia doveva affrontare, il “disgusto” per la mancanza di mascherine e gel disinfettante e la “tristezza” delle persone che ogni giorno facevano i conti con il bollettino dei nuovi casi/morti che gettava sempre più nello sconforto.

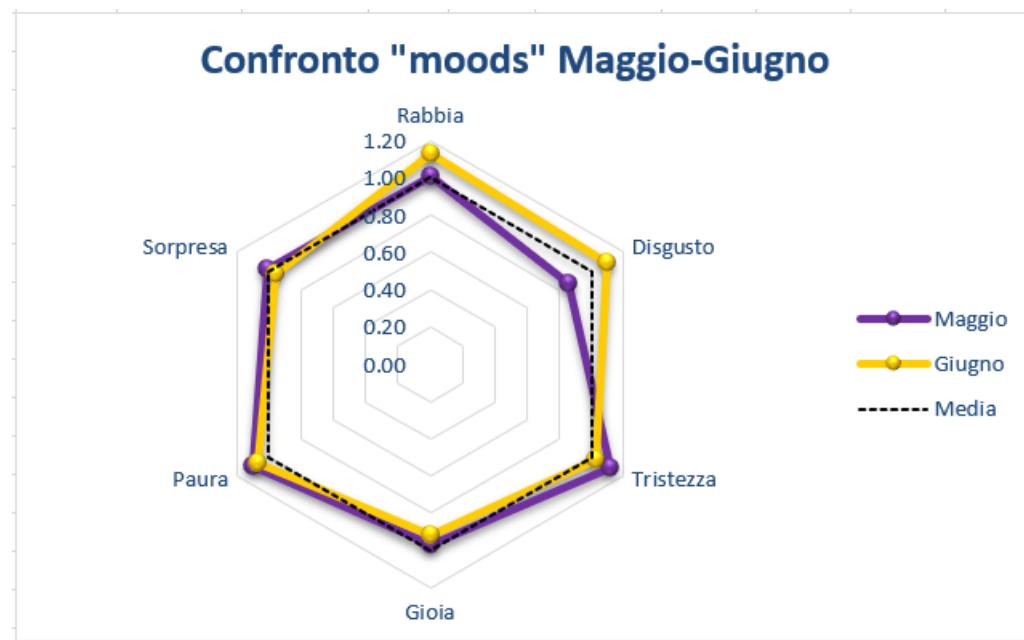
Anche la situazione delle “gioia” conferma la difficoltà della situazione che si era

creata, con un importante decremento tra marzo e aprile (primi effetti ed impatto del virus).

Confrontando con la media vediamo che, grosso modo, “*rabbia*” e “*tristezza*” si aggirano intorno a quei livelli, ciò significa che dopo l’incremento nel primo mese queste emozioni hanno mantenuto questa tendenza, non ritornando ai valori iniziali. La “*paura*”, a differenza di quanto detto finora, supera il livello medio, effetto ragionevole poiché nei primi mesi di pandemia, essendo la prima volta che si affrontava una situazione del genere, vi era un po’ di “*paura*” in ogni persona.

Per quanto riguarda invece il periodo marzo-giugno:

Grafico 11. Confronto “moods” Maggio-Giugno



Fonte: Elaborazione personale, Excel

A giugno vediamo che “*tristezza*” e “*paura*” iniziano un po’ a diminuire dopo un aumento nei mesi *core* della pandemia. Tuttavia, emozioni come “*rabbia*” e “*disgusto*”

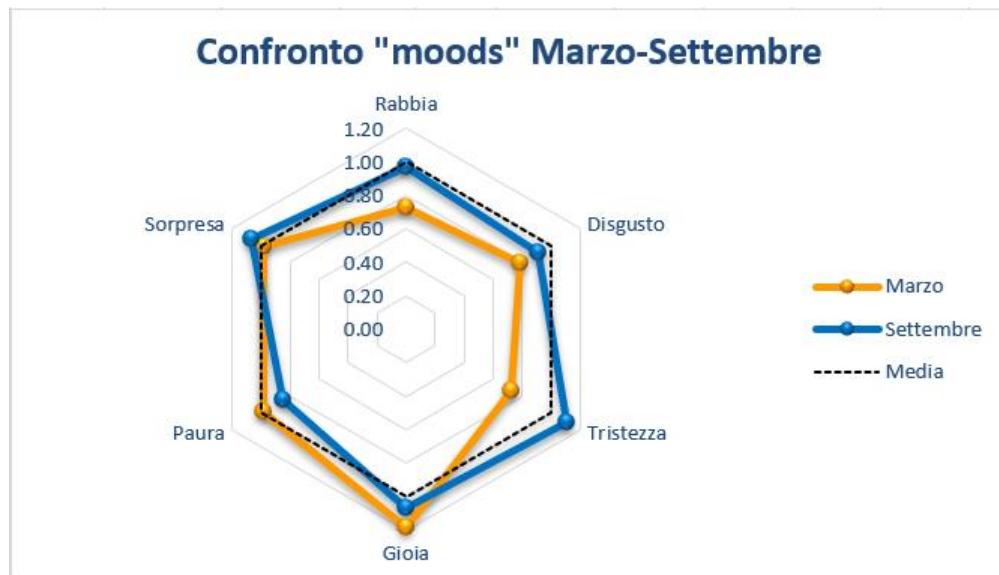
mostrano addirittura un incremento tra maggio e giugno, nonostante in questi mesi le restrizioni sono state allentate e si andava incontro verso a delle graduali aperture.

Perché allora questo?

Un motivo potrebbe essere l'uscita dell'*App Immuni* il 15 giugno che ha fatto molto scalpore, scatenando teorie complottistiche nelle persone più scettiche; infatti, ciò che aumenta in modo più netto è proprio il “*disgusto*”, più che la rabbia, andando oltre la media (nella tabella precedente avevamo visto che in questi due mesi vi era un forte incremento di queste due emozioni).

Nonostante ciò, vi inizia ad essere una ripresa “emozionale” a causa proprio delle riaperture e del diminuire, giorno dopo giorno, dei casi *Covid-19*. Una forte ripresa c’è stata a settembre, un mese di “rinascita”:

Grafico 12. Confronto "moods" Marzo-Settembre

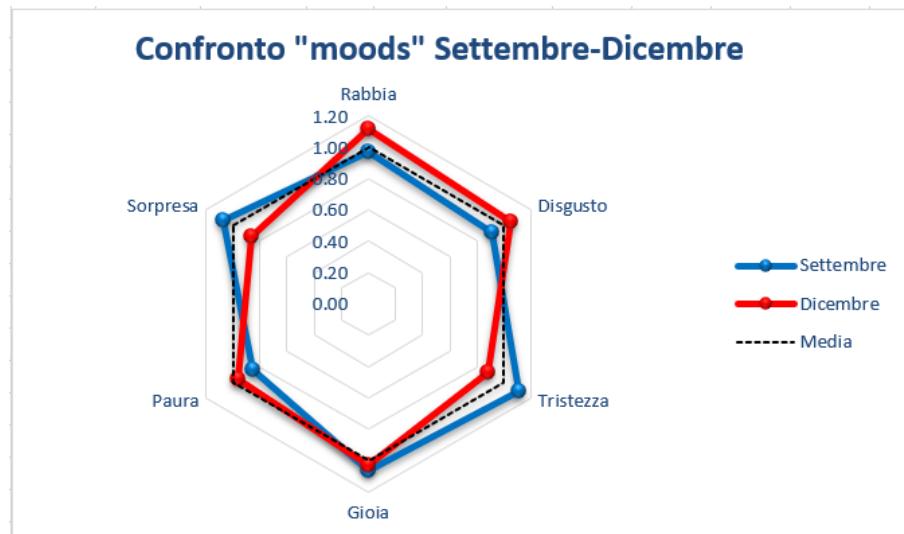


Fonte: Elaborazione personale, Excel

A settembre si vede come “*disgusto*” e “*paura*” scendono sotto la media, segni proprio di una rinascita emozionale, la “*rabbia*” invece rimane costante intorno la media. A conferma di ciò abbiamo anche l’aumento della “*sorpresa*”, che ricordiamo essere un’emozione positiva, che supera addirittura i livelli di marzo, ossia pre-pandemia. I livelli di marzo delle altre emozioni, invece, non vengono raggiunti, nonostante la ripresa non si può affermare di essere tornati all’inizio, vediamo come la “*gioia*” di settembre rimane decisamente più bassa, così come “*rabbia*” e “*disgusto*” a marzo siano nettamente inferiori a settembre; caso rilevante è quello della “*tristezza*” che non vede un calo in questo periodo ma addirittura un aumento superando la soglia media, questo dovuto probabilmente alla forte depressione che il *Covid-19* ha portato con sé.

Infine, è stato necessario confrontare anche il periodo settembre-dicembre, ossia il mese della rinascita con il mese *core* della seconda ondata che ha portato ad ulteriori restrizioni, compromettendo il clima natalizio:

Grafico 13. Confronto "moods" Settembre-Dicembre



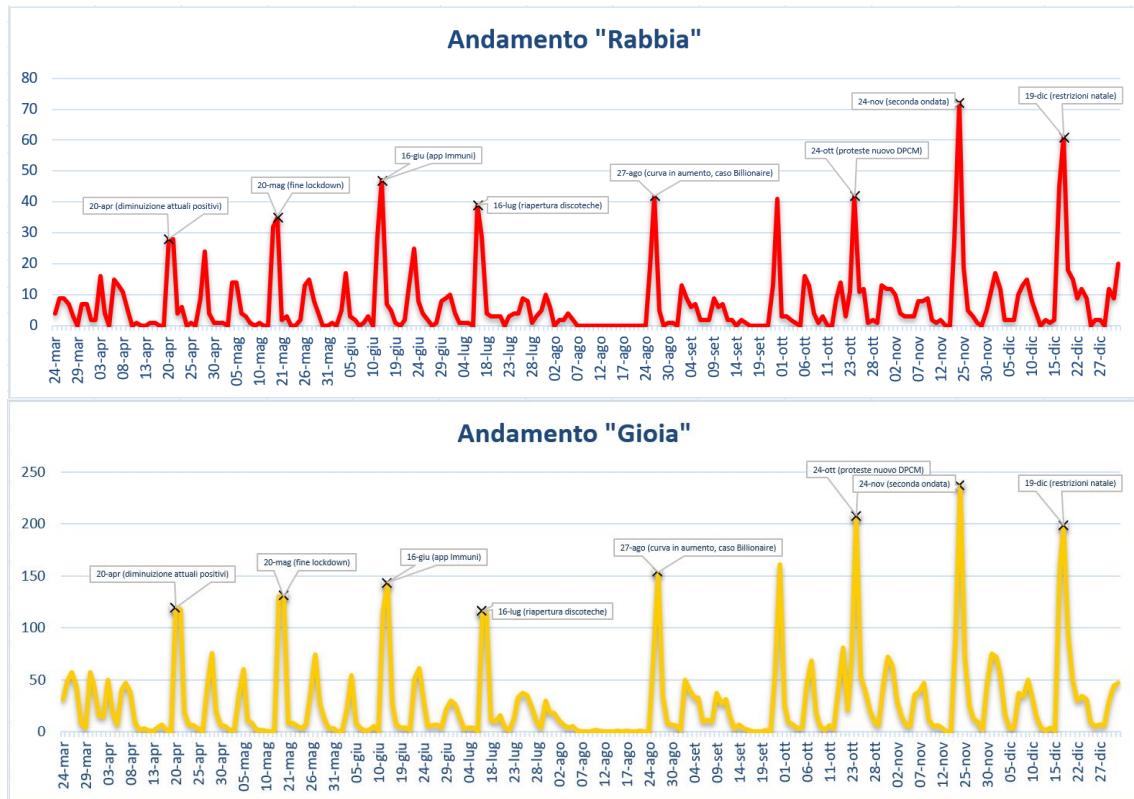
Fonte: Elaborazione personale, Excel

La prima cosa che viene fuori è il grande aumento di “*rabbia*” e “*disgusto*” giustificato dalle restrizioni che il 18 dicembre hanno compromesso il clima natalizio 2020, le quali evitavano gli spostamenti tra comuni e la visita solo ad un numero limitato di conoscenti. Questo tema della “*rabbia*” è stato dominante in questo mese, probabilmente la “*tristezza*” che si aveva prima, la profonda depressione si è proprio trasformata in “*rabbia*” a causa di questa seconda ondata e delle nuove restrizioni (oltre alla divisione delle regioni per colore), e per questo sembrerebbe essere al di sotto addirittura della media. Si vede poi come anche la “*sorpresa*” (emozione che esprime qualcosa di positivo) diminuisce scendendo al di sotto della media, e la “*paura*” che ritorna di moda a causa dei numerosi contagi.

Bisogna sottolineare, però, che questo periodo ha prodotto un caleidoscopio di emozioni, di conseguenza ogni mese non ha visto solo l'avanzare di un'emozione in particolare, ma di tutte; in un mese in cui c'è stata forte “*rabbia*” vi sono comunque segni di “*gioia*”, questo perché oltre alle condizioni esterne c'è la banca che ha continuato a fare il suo lavoro, le emozioni grosso modo si sono distribuite sempre in modo simile, ciò che è stato fatto era cercare di individuare quei *pattern* un po' più evidenti che hanno caratterizzato determinati momenti.

Per capire meglio ciò bisogna guardare il successivo grafico:

Grafico 14. Andamento "Rabbia" e "Gioia"



Fonte: Elaborazione personale, Excel

Si può facilmente vedere come i picchi spesso sono derivati da eventi specifici, ciò non vuol dire però che ad un determinato evento bisogna associare unicamente un'emozione; nel primo dei due grafici si ha l'andamento della “*rabbia*”, nonostante ciò, si vedono dei picchi anche per situazioni oggettivamente positive (es. fine *lockdown*), questo perché ogni situazione è stata diversa per ciascuna persona, ognuno reagiva in maniera differente, non è detto che se una situazione è scomoda deve necessariamente esprimere “*rabbia*”.

Il 15 giugno, per esempio, vi è un forte picco a causa dell’uscita dell’*App Immuni*,

guardando solo questo grafico si potrebbe dire che ciò ha causato un innalzamento del “*mood*” “*rabbia*”, ma è vero solo in parte, perché vedendo anche il grafico dell’andamento della “*gioia*” si può trovare un picco in questa stessa data, perché se alcune persone erano contrarie all’*App*, creando teorie complottiste, altre erano del tutto felici perché si cercava di contenere i contagi e di rintracciarli il più velocemente possibile. Stessa cosa vale ragionando sulla banca, in uno stesso periodo vi sono stati fattori di forte detrazione, come “*assenza di sportelli*”, “*problemi legati all’app e sito*”, ma anche fattori che si sono rivelati decisamente positivi come “*disponibilità/cortesia*” o “*professionalità*”, a conferma del fatto che in particolari date si stressano un po’ tutte le emozioni.

Questo discorso vale per tutte le emozioni. Inoltre, a tutto ciò bisogna aggiungere che non tutta la “*gioia*”, così come la “*rabbia*”, deriva da fattori esterni, anzi, il *trend* principale è dovuto proprio dalla banca stessa in base alle sue *performance*, che naturalmente è stato influenzato in parte dalle situazioni difficili che vi erano fuori.

In sintesi, analizzando e confrontando un po’ tutti i grafici visti in questo paragrafo, si potrebbe dire che il *Covid* ha portato un aumento delle emozioni negative fin da subito. A causa poi di alcuni eventi, “*rabbia*” e “*disgusto*” sono venute fuori specialmente intorno giugno e dicembre, mentre nel periodo tra questi due mesi sembrerebbe esserci una sorta di fase di depressione dovuta all’aumento di “*tristezza*” e “*paura*”.

Verso settembre le emozioni positive sono tornate a buoni livelli a causa della diminuzione dei casi, senza però riuscire a ritornare ai livelli iniziali.

Una volta visto l'andamento generale delle emozioni ci si è chiesto se alcune emozioni erano presenti in particolare segmenti di persone oppure no, perché lo scopo di questo elaborato è di capire, fondamentalmente, se ci sono relazioni tra le emozioni di una persona, i “*moods*” e le caratteristiche della stessa, in modo da sfruttare ciò in futuro per cercare di prevedere lo stato d'animo e capire *ex-ante* come trattare questi clienti (es. se so che una persona ha “*paura*” provo a rassicurarla in qualche modo).

Alcune delle caratteristiche più importanti in questo caso, ma generalmente essenziali in ogni analisi, sono sicuramente l'*età*, il *livello di scolarizzazione* e il *reddito* di una persona.

In questo *dataset* non abbiamo in modo diretto queste variabili, ma bisogna considerare che *personas* è una segmentazione fatta proprio su queste tre caratteristiche; quindi, risulta facile ottenere queste informazioni per ogni osservazione partendo dalla classe di *personas* associata (es. se un'osservazione appartiene a *Teresa* sappiamo che ha una scolarizzazione bassa, un'età superiore ai 60 anni e un reddito alto).

Per quanto riguardo l'età si ha questa situazione:

*Tabella 10.*

Classe età	Rabbia	Disgusto	Tristezza	Gioia	Paura	Sorpresa	Totale
<b>19-25</b>	0.10	0.11	0.07	<b>0.45</b>	0.17	0.10	1.00
<b>26-45</b>	<b>0.11</b>	<b>0.12</b>	0.07	0.41	0.19	<b>0.10</b>	1.00
<b>46-60</b>	0.10	<b>0.11</b>	0.07	0.42	0.19	<b>0.10</b>	1.00
<b>Over 60</b>	<b>0.11</b>	0.09	0.06	<b>0.44</b>	0.20	<b>0.10</b>	1.00
<b>Totale</b>	0.10	0.11	0.07	0.43	0.19	0.10	1.00

*Fonte: Elaborazione personale, Excel*

Dalla tabella sembrerebbe che con l'aumentare dell'età aumenti anche la “*paura*”; infatti, abbiamo la percentuale più bassa proprio nei giovani, la classe 19-25; per quanto riguarda “*rabbia*” e “*disgusto*” si hanno delle situazioni abbastanza simili (come già visto finora) che vedono sempre la classe dei giovani quella meno colpita.

Discorso diverso per la “*gioia*”, che non fa altro che confermare quanto detto, ossia la maggiore serenità delle persone giovani e degli over 60 a discapito della clientela compresa tra i 25 e 60 anni (classe di età che comprendeva, come visto precedentemente, una maggiore detrazione).

La variabile che probabilmente dice di più, facendo una distinzione più netta è il livello di scolarizzazione che mostra questi risultati:

*Tabella 11.*

Scolarizzazione	Rabbia	Disgusto	Tristezza	Gioia	Paura	Sorpresa	Totale
<b>Alta</b>	0.11	0.12	0.07	0.38	0.22	0.10	1.00
<b>Bassa</b>	0.10	0.10	0.07	0.46	0.18	0.10	1.00
<b>Totale</b>	0.10	0.11	0.07	0.42	0.20	0.10	1.00

*Fonte: Elaborazione personale, Excel*

In questo caso si vede una netta distinzione tra “alta scolarizzazione” e “bassa scolarizzazione” per quanto riguarda la “gioia” e la “paura”; i gioiosi sembrerebbero quelli meno istruiti e, di conseguenza, le persone più impaurite quelle più istruite.

All’aumentare della scolarizzazione sembra aumentare anche la “rabbia”, il “disgusto” e la “tristezza”, seppur in modo meno evidente e netto.

Infine, vi è il reddito, suddiviso in “alto” e “basso”:

*Tabella 12.*

Reddito	Rabbia	Disgusto	Tristezza	Gioia	Paura	Sorpresa	Totale
<b>Alto</b>	0.11	0.11	0.07	0.42	0.20	0.10	1.00
<b>Basso</b>	0.10	0.11	0.07	0.44	0.18	0.10	1.00
<b>Totale</b>	0.10	0.11	0.07	0.43	0.19	0.10	1.00

*Fonte: Elaborazione personale, Excel*

Anche qui si mostra un andamento simile a quello visto sopra, solamente che la

variabile in questione adesso è il reddito; all'aumentare del reddito aumenterebbero leggermente “*paura*” e “*rabbia*”, mentre al suo diminuire aumenterebbe la “*gioia*” delle persone.

In questo caso il proverbio “i soldi non fanno la felicità” è d’obbligo.

Per una conferma sono stati fatti anche i relativi “*grafici a radar*”:



*Fonte: Elaborazione personale, Excel*

Il *grafico a radar* conferma in modo più chiaro quanto detto sopra, gli over 60 sono la classe con meno “*disgusto*” e “*tristezza*” poiché ricordiamo che parliamo di una classe di pensionati (probabili nonni) e magari il dover stare chiusi in casa ha avvicinato nipoti e parenti; tuttavia, sembrerebbero essere quelli con più “*paura*”, dovuto magari al fatto che il *Covid* è più pericoloso per le persone con l’età avanzata. La fascia con più

“rabbia”, “tristezza” e “disgusto” solo quelli tra i 25 e 60 anni, persone spesso al centro della loro carriera lavorativa.

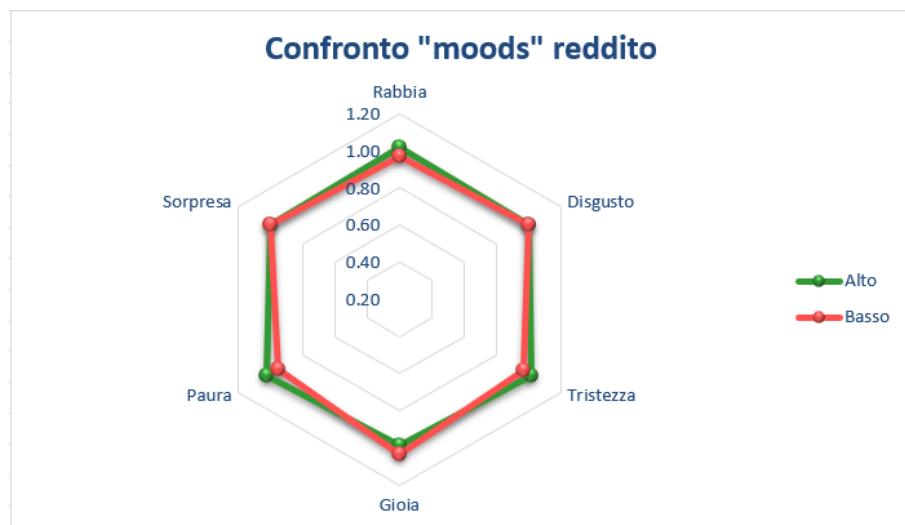
Grafico 16. Confronto "moods" in base alla scolarizzazione



Fonte: Elaborazione personale, Excel

Anche qui il *grafico a radar* conferma quanto ipotizzato sopra, ossia che le persone più felici sono quelle meno scolarizzate, mentre le persone scolarizzate sono le più stressate (*mix di “rabbia”, “tristezza”, “paura” e “disgusto”*).

Grafico 17. Confronto "moods" in base al reddito



Fonte: Elaborazione personale, Excel

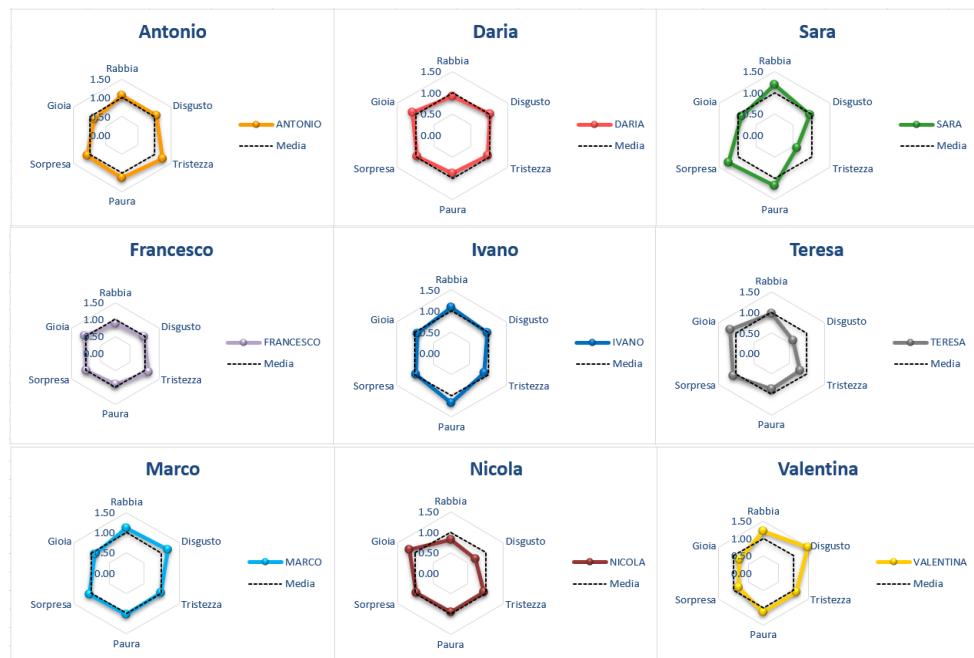
Anche qui, come visto sopra, sembrerebbe (in modo più leggero) che un reddito più alto porti un maggiore stress emozionale, poiché si ha la stessa situazione che si ha con la scolarizzazione.

Facendo una combinazione dei tre grafici è possibile creare una sorta di “*profilo*” del cliente che ha subito un maggiore stress emozionale, ossia una persona con:

- Età compresa tra 25 e 60 anni
- Alta scolarizzazione
- Alto reddito

Per avere una conferma di ciò si può andare ad analizzare la distribuzione dei “*moods*” per ciascuna *personas*:

Grafico 18. "Moods" di ogni Personas



Fonte: Elaborazione personale, Excel

Dai *grafici a radar* può essere confermato quanto detto fino a questo punto, partendo dai segmenti scolarizzati si può notare che tendono ad avere più “*rabbia*” e “*disgusto*” rispetto alla media, in particolar modo la fascia 26-60 con *Valentina*, seguita da *Marco* e *Antonio*.

Gli over 60, al contrario, sono quelli con meno “*moods*” negativi, si prenda ad esempio *Nicola*, i cui livelli di “*rabbia*” e “*disgusto*” sono i più bassi, seguito da *Teresa*, la quale sembrerebbe avere della “*rabbia*” in più, magari dovuto al fatto di essere un segmento con reddito alto; inoltre, sono anche i segmenti con più “*gioia*”, seguiti da *Daria*, una conferma che i segmenti meno istruiti sembrerebbero i più felici. Il segmento *Sara*, fondamentalmente i giovani, è quello che ha avuto meno “*tristezza*”, come si poteva intuire.

Il segmento che sembra aver subito di più questo periodo difficile, pieno di un *mix* di emozioni, di stress, è *Valentina*, e lo possiamo notare dal fatto che tutte le emozioni negative (“*paura*”, “*rabbia*”, “*disgusto*” e “*tristezza*”) sono superiori alla media, mentre le emozioni positive inferiori.

Per poter avere una conferma su quanto ipotizzato è stato aggiornato il *dataset* iniziale con altre variabili, tra cui la professione. La situazione che è venuta fuori è la seguente:

*Tabella 53.*

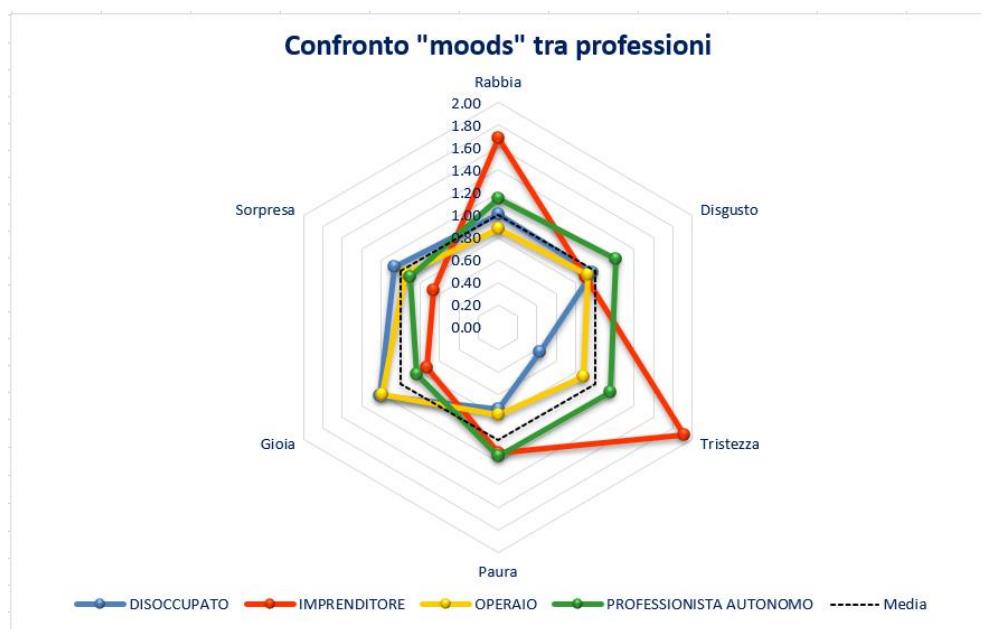
Professioni	Rabbia	Disgusto	Tristezza	Paura	Gioia	Sorpresa	Totale
<b>ALTRO</b>	<b>0.14</b>	<b>0.15</b>	0.15	0.22	0.25	0.09	1.00
<b>CASALINGA</b>	0.08	0.10	0.08	0.18	0.47	0.08	1.00
<b>DISOCCUPATO</b>	0.11	0.11	0.03	0.14	0.51	0.11	1.00
<b>IMPIEGATO</b>	0.11	0.13	0.07	0.21	0.39	0.09	1.00
<b>IMPIEGATO PA</b>	0.12	0.08	0.09	0.16	0.46	0.10	1.00
<b>IMPRENDITORE</b>	0.18	0.10	0.14	0.22	0.31	0.06	1.00
<b>INSEGNANTE</b>	0.10	0.11	0.06	0.22	0.39	0.12	1.00
<b>MANAGER/DIRIGENTE</b>	0.12	0.14	0.10	0.20	0.34	0.11	1.00
<b>MILITARE</b>	0.08	0.09	0.08	0.21	0.41	0.12	1.00
<b>OPERAIO</b>	0.09	0.10	0.06	0.15	0.50	0.09	1.00
<b>PENSIONATO</b>	0.10	0.09	0.06	0.20	0.45	0.10	1.00
<b>PROFESSIONISTA AUTONOMO</b>	0.12	0.13	0.08	0.23	0.35	0.09	1.00
<b>REDDITIERO</b>	0.17	0.23	0.06	0.17	0.31	0.06	1.00
<b>STUDENTE</b>	0.11	0.12	0.04	0.18	0.42	0.13	1.00
<b>Totale</b>	0.10	0.11	0.07	0.20	0.42	0.10	1.00

*Fonte: Elaborazione personale, Excel*

La tabella mostra il *mix* di emozioni di ciascun cluster di professioni (le professioni sono state raggruppate); ciò che ne viene fuori è molto interessante e tende a confermare quanto ipotizzato finora. Come prima cosa si va a vedere la “gioia”, emozione che supera la media in professioni umili come “operaio”, “casalinga” e persone che al momento non lavoravano, come “disoccupati” e “pensionati”. Per quanto riguarda queste ultime due professioni la risposta potrebbe essere semplice, dovuta semplicemente al fatto che tali clienti non rischiavano di perdere un posto di lavoro, situazione che rende tutt’altro che felici. Interessante è la differenza tra “impiegati” e “impiegati della PA”, poiché hanno risposte emozionali completamente differenti, livelli bassi di “gioia” per il primo e alti per il secondo; questo perché gli “impiegati della PA”

potevano perdere più difficilmente il lavoro, soprattutto in una situazione come quella vista. Una conferma la possiamo trovare analizzando professioni di alto rango, come “imprenditori”, “professionisti autonomi” e “manager/dirigenti”; la situazione qui è completamente diversa, si possono notare livelli di “gioia” tra i più bassi, e allo stesso tempo livelli di emozioni negative, quali “rabbia”, “disgusto”, “paura” e “tristezza”, tra i più elevati, a conferma del fatto che tali cluster di professioni sono tra quelli che hanno risentito maggiormente del periodo difficile. La risposta anche qui sembrerebbe abbastanza semplice; la pandemia *Covid-19*, con il successivo *lockdown*, ha reso la vita difficile a tantissime attività commerciali (e non), portando al fallimento di molte e, di conseguenza, “imprenditori” e professioni di rango simile si trovavano in una costante lotta per non vedere fallire la propria attività. Confrontando nello specifico queste professioni in un grafico a radar, dopo aver calcolato la *lift*, il risultato è il seguente:

Grafico 19. Confronto "moods" tra professioni, grafico a radar



Fonte: Elaborazione personale, Excel

In primis si può notare l'elevato livello di “*tristezza*” e “*rabbia*” da parte degli “*imprenditori*” e il basso livello di “*gioia*” e “*sorpresa*”, a conferma di tutto ciò che è stato detto; per quanto riguardo i “*professioni autonomi*”, non si hanno gli stessi livelli degli “*imprenditori*” ma sono comunque molto elevati, al di sopra della media, per le emozioni negative, e al di sotto, per le emozioni positive. Per i “*disoccupati*” il livello di “*tristezza*” è bassissimo, insieme a quello di “*paura*” ed è coerente con quanto sostenuto perché si tratta di una categoria di persone che non rischiano di perdere nessun lavoro.

Mettendo insieme quanto visto fin qui, tra professioni, reddito, livello di scolarizzazione ed età, si potrebbe azzardare nel dire che coloro che hanno risentito di più nel periodo *Covid 2020* sono state persone con tali caratteristiche:

- Età compresa tra 25-60
- Alta scolarizzazione
- Alto reddito
- Professioni di alto rango.

## 2.5 La felicità incide sulla raccomandabilità?

Prima di poter arrivare a costruire un modello che provi a determinare il “*mood*” di una persona a priori, in base a determinate caratteristiche, è stato necessario fare una prova del nove che confermasse quanto di buono detto finora.

Intuitivamente verrebbe da pensare che una persona arrabbiata potrebbe essere un problema per la banca, iniziando a recensirla in modo negativo e infine a decidere di lasciarla e viceversa per quanto riguarda i clienti felici. Per questo motivo è stato necessario andare a vedere la composizione dei voti *NPS* rispetto ai “*moods*” dei clienti, per capire se quanto fatto finora ha senso o se magari chi ha dato voti negativi è stato classificato come “*gioia*”.

*Tabella 14.*

Voto NPS	Rabbia	Disgusto	Tristezza	Paura	Sorpresa	Gioia
0	0.33	0.41	0.32	0.22	0.18	0.08
1	0.08	0.08	0.08	0.05	0.04	0.02
2	0.07	0.07	0.06	0.06	0.04	0.02
3	0.07	0.07	0.08	0.06	0.06	0.02
4	0.07	0.05	0.06	0.06	0.03	0.02
5	0.10	0.08	0.09	0.10	0.06	0.04
6	0.08	0.06	0.07	0.08	0.05	0.03
7	0.08	0.07	0.08	0.09	0.05	0.07
8	0.06	0.06	0.08	0.10	0.07	0.14
9	0.03	0.03	0.04	0.07	0.12	0.20
10	0.03	0.02	0.04	0.11	0.30	0.37

*Fonte: Elaborazione personale, Excel*

La tabella mostra come ogni emozione si distribuisce per ciascun voto *NPS* (con un totale di colonna pari a 1). Sembrerebbe che le emozioni negative, quali “*rabbia*”, “*paura*”, “*disgusto*” e “*tristezza*” abbiano una polarizzazione maggiormente nei voti di

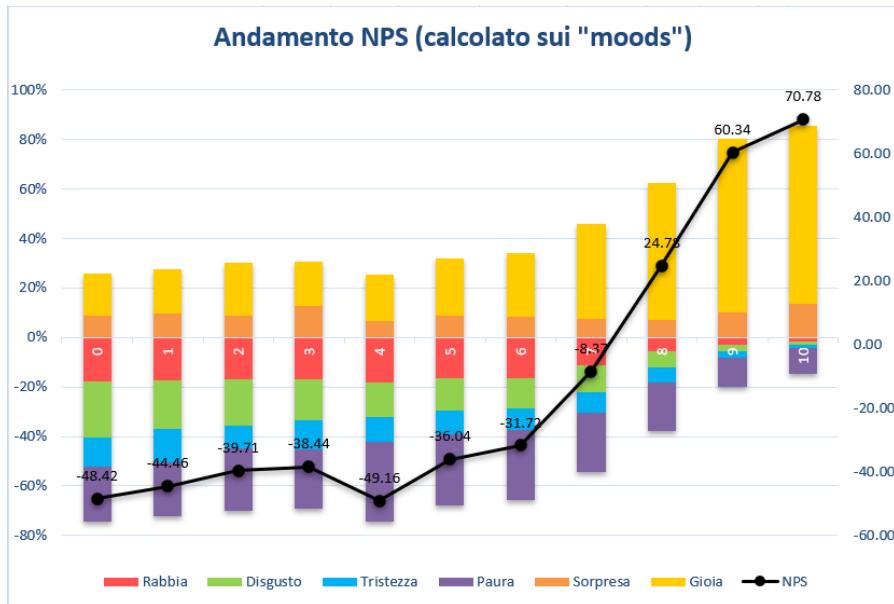
detrazione (voti 0-6), in particolare nel voto 0, e ciò è coerente con quanto detto e fatto finora poiché clienti arrabbiati tenderebbero a valutare negativamente la banca e di conseguenza, dall'altra parte della medaglia, abbiamo che emozioni positive, quali “gioia” e “sorpresa” si concentrano maggiormente nei voti alti (9-10).

Una piccola precisione va fatta sulle emozioni positive che sembrerebbero avere un'alta percentuale anche per i clienti che hanno dato il voto 0, il che non ha molto senso, ed infatti questo effetto, questa sovrastima di “gioia” e “sorpresa”, è probabilmente dovuto all'algoritmo usato per l'estrazione dei “mood” dai *Verbatim*, il quale non è affatto ottimizzato per la lingua italiana, viene per le prime volte utilizzato ed è in continuo sviluppo, di conseguenza ci può stare una lieve distorsione come questa.

Nonostante ciò, l'algoritmo sembrerebbe funzionare piuttosto bene perché i risultati che mostra questa tabella sono molto incoraggianti e pertinenti al discorso fatto finora.

Una ulteriore prova è stata fatta calcolando un *NPS* considerando come “*detrattori*” le emozioni negative e come “*promotori*” le emozioni positive:

Grafico 20. Andamento NPS (calcolato sulle emozioni)



Fonte: Elaborazione personale, Excel

Il grafico mostra molto bene l'andamento dell'*NPS* per ogni classe di voto.

Si parte da un *NPS* estremamente negativo per le classi di voto tipiche dei detrattori (0-6), con un leggero incremento che è coerente poiché aumentando il voto, di conseguenza, sembrerebbe aumentare anche l'*NPS*, ciò vuol dire che aumentando il voto sembrerebbero diminuire le emozioni negative e aumentare quelle positive.

Nella classe di voto 0-6 le emozioni negative, infatti, prevalgono sulle positive fino ad arrivare nella classe 7-8 (i passivi) dove si inizia a vedere un netto aumento dell'*NPS* perché iniziano a diminuire le emozioni negative a favore delle positive.

L'inversione di marcia si ha tra i voti 7-8 poiché l'*NPS* da negativo inizia a diventare positivo, segno che le emozioni positive iniziano a prevalere su quelle negative,

ricordando però che la classe 7-8 è quella dei passivi, ossia clienti indifferenti, anche se i passivi dal voto 7 sembrerebbero più vicini ai detrattori (voti 0-6), mentre i passivi dal voto 8 più vicini ai promotori (voti 9-10) in termini di emozioni.

Arrivati nella classe dei promotori (voti 9-10) si vede come l'*NPS* cresce in modo vertiginoso, segno che le emozioni positive superano nettamente le negative, ed è coerente considerando che questa classe comprende i clienti che hanno dato i voti più alti, ossia sono i clienti più soddisfatti.

Per avere un’ulteriore conferma dell’importanza delle emozioni, per prevedere i clienti più propensi ad abbandonare la banca, è stata analizzata anche la variabile “*Fedeltà*” che comprende tre modalità:

- *Intenzione di abbandonare la banca nei prossimi 12 mesi;*
- *Intenzione di valutare le offerte di altre banche nei prossimi 12 mesi;*
- *Intenzione di proseguire l’esperienza con la banca;*

I risultati sono i seguenti:

*Tabella 15.*

<b>Fedeltà</b>	<b>Rabbia</b>	<b>Disgusto</b>	<b>Tristezza</b>	<b>Paura</b>	<b>Sorpresa</b>	<b>Gioia</b>
<b>Chiudere rapporti</b>	0.24	0.36	0.27	0.18	0.16	0.08
<b>Valutazione altre offerte</b>	0.46	0.43	0.43	0.40	0.27	0.17
<b>Proseguire esperienza</b>	0.30	0.22	0.30	0.42	0.57	0.75

*Fonte: Elaborazione personale, Excel*

Anche qui la situazione sembra coerente poiché “*rabbia*”, “*disgusto*” e “*tristezza*”

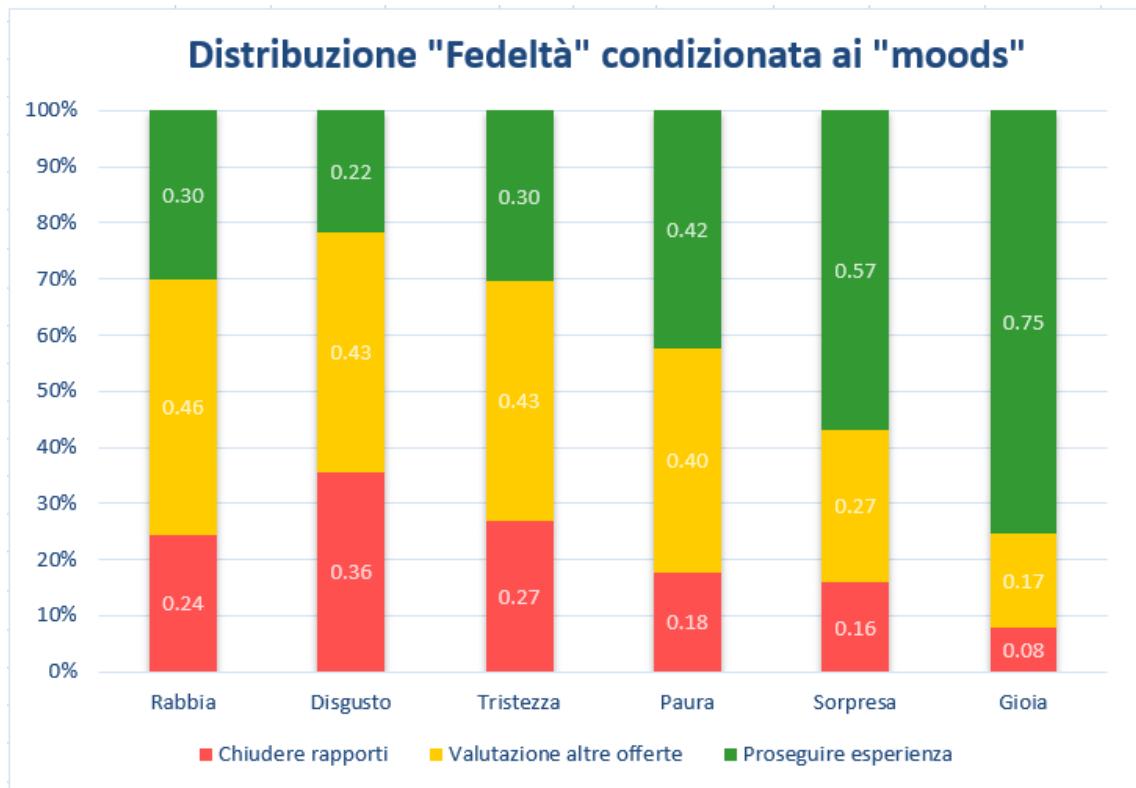
hanno una percentuale molto alta di clientela che intende lasciare la banca, o che comunque intende valutare altre offerte. Si nota come, probabilmente, il “*mood*” più forte, in senso negativo, è il “*disgusto*”, dove più di 1 persona su 3 intende lasciare la banca e solamente 1 su 5 intende proseguire l’esperienza con la banca. Rispetto alla situazione vista prima, la “*paura*” non è un’emozione che tende a far perdere clientela fortemente, si potrebbe dire che tende a creare dei dubbi riguardo la banca, costringendo le persone a guardarsi intorno; nonostante tutto 4 persone su 10 che riscontrano “*paura*” decidono comunque di restare con la banca.

Infine, “*gioia*” e “*sorpresa*”, le emozioni positive, effettivamente mostrano una situazione in cui il cliente preferisce rimanere con la banca, perché molto probabilmente soddisfatto della stessa; questa situazione, ovviamente è molto più evidente con la “*gioia*” piuttosto che con la “*sorpresa*”, la quale è un’emozione un po’ più complicata da gestire e decifrare.

Nonostante ciò, si potrebbe trovare il “*pelo nell’uovo*”, ossia quei clienti che nonostante esprimano “*gioia*” e “*sorpresa*” intendono chiudere i rapporti con la banca e le cui percentuali possono essere abbastanza rilevanti; anche qui, questo effetto è probabilmente dovuto alle limitazioni del modello usato per estrarre i “*mood*” dai *Verbatim* e fondamentalmente si tratta delle stesse persone che nella tabella precedente hanno dato votazione 0 ma che sono state identificate con la “*gioia*”, ovviamente in modo errato.

Si può avere una panoramica migliore dal seguente grafico:

*Grafico 21. Distribuzione "Fedeltà" condizionata ai "moods"*



## 2.6 Un indicatore di attualità

Prima di arrivare alla parte finale, ossia il modello, è stato proposto un indicatore di “*attualità*” che, in qualche modo, cerca di racchiudere tutte le informazioni necessarie.

Se si ripensa al periodo *Covid-19*, una delle cose che sicuramente viene può venire in mente è l'introduzione dei colori per le regioni per differenziarle in base allo stato di emergenza; vi sono state diverse suddivisioni del territorio italiano e tutte avevano in comune la “*zona rossa*” per identificare una particolare regione in cui la situazione era piuttosto grave, con numerosi contagi etc.

L'indicatore, che di seguito verrà proposto, cerca di racchiudere proprio questo tema delle restrizioni e non solo, perché comprende anche un tema riguardo il “*tempo*” e lo “*spazio*”.

In che senso?

Avendo a disposizione la data di ogni intervista fatta, è stato possibile associare a ciascuno, considerando anche la regione di provenienza, il colore della stessa al momento dell'intervista.

Es. Ad un cliente di “*Toscana e Umbria*” che ha risposto al questionario l’11 novembre è stata assegnata la “*zona arancione*”, colore della regione in quel momento.

Bisogna precisare, però, che il sistema delle regioni “colorate” è entrato in vigore attraverso apposito *DPCM* il 6 novembre 2020 (tre diverse zone, *rossa*, *arancione* e *gialla*), quindi a tutte le interviste precedenti è stata assegnata una “*zona di restrizione*”

in base alle similarità e limitazioni che si sono avute poi con l'introduzione il 6 novembre.

Il criterio che è stato utilizzato è il seguente:

- Alle interviste di marzo è stata assegnata la “*zona gialla*” perché, nonostante ci fosse già il *lockdown*, è un po' il punto di partenza in cui non erano molto presenti gli effetti del virus;
- Dal 1° aprile fino al 22 giugno, quindi fino all'inizio della fase 3 e, in particolare, al giorno con minori positivi dall'inizio della pandemia, è stata assegnata la “*zona rossa*”;
- Dal 23 giugno al 2 luglio, date le varie aperture, è stata assegnata la “*zona arancione*”;
- Dal 3 luglio al 17 ottobre è stata assegnata la “*zona gialla*”;
- Dal 18 ottobre fino al 5 novembre, dato l'aumento dei contagi a causa della seconda ondata, è stata assegnata la “*zona arancione*”;
- Dal 6 novembre in poi sono state prese in considerazioni le colorazioni assegnate dal Governo.

Dopo aver creato questo “*indicatore di restrizione*”<sup>33</sup> è stata necessaria un’analisi della correlazione, per vedere se il comportamento dello stesso è coerente con le emozioni dei clienti. Avendo, però, tutta variabili dicotomiche (0-1) è stata effettuata la “*correlazione tetracorica*” nel seguente modo:

*Figura 5.*

	<i>modalità 1</i>		<i>modalità 2</i>	<i>tot.</i> $a+b$ $c+d$ $n=a+b+c+d$
	<i>modalità A</i> $a$ $c$ $a+c$	<i>modalità B</i> $b$ $d$ $b+d$		

*Fonte: Internet*

Il grado di associazione tra le due variabili dicotomiche si saggia attraverso l’indice  $\varphi$  *tetracorico*:

$$\varphi = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$$

Una volta calcolate le correlazioni attraverso tale indice sono stati ottenuti i seguenti risultati:

*Tabella 16.*

	Rabbia		Disgusto
Giallo	0.008	Rosso	0.020
Rosso	0.001	Arancione	0.007
Arancione	-0.010	Giallo	-0.026
	Gioia		Sorpresa
Giallo	0.017	Giallo	0.011
Arancione	-0.008	Arancione	-0.001
Rosso	-0.011	Rosso	-0.011
	Tristezza		Paura
Giallo	0.004	Rosso	0.021
Rosso	-0.001	Arancione	-0.002
Arancione	-0.004	Giallo	-0.018

*Fonte: Elaborazione personale, Excel*

<sup>33</sup> Non è stato possibile associare una zona a tutti i clienti poiché alcuni non avevano la regione, essendo clienti *internet*

Prima di analizzare la tabella precedente, va precisato che non interessa l'intensità della correlazione, ma l'ordine e, in particolar modo, il segno.

Partendo dalle emozioni positive (*gioia* e *sorpresa*) sembra che la “*zona gialla*” sia quella che abbia portato più “*good vibes*” e lo notiamo sia dal segno positivo che dall'ordine, a seguire rispettivamente “*zona arancione*” e “*zona rossa*”.

Per le emozioni negative vediamo una situazione un po' invertita, con “*zona rossa*” al primo posto e con segno positivo (quindi correlata positivamente) sia per quanto riguarda il “*disgusto*” che la “*paura*”; una differenza tra queste due emozioni la si può trovare nella “*zona arancione*”, poiché sembrerebbe che codesta sia correlata positivamente con il “*disgusto*”, ma negativamente con la “*paura*”. Per “*rabbia*” e “*tristezza*” una situazione un po' ambigua è la prima posizione della “*zona gialla*”.

Per la “*tristezza*”, inoltre, vi è una correlazione negativa sia con la “*zona rossa*” che con la “*zona arancione*”, il che può essere dovuto ad una trasformazione dell'emozione da “*tristezza*” a “*paura*” con l'aumentare delle restrizioni.

Nel complesso l'indicatore sembra abbastanza coerente con la situazione delle emozioni.

Per una ulteriore conferma è stata implementata un'analisi delle modalità caratterizzanti per vedere se alcune emozioni potevano essere associate ad una determinata “zona”:

*Tabella 17.*

Zona Rossa					Zona Arancione					Zona Gialla				
	No	Si	Sign.	Lift		No	Si	Sign.	Lift		No	Si	Sign.	Lift
<b>Disgusto</b>	0.126	0.141	,051	1.113	<b>Disgusto</b>	0.130	0.136	,473	1.048	<b>Sorpresa</b>	0.116	0.123	,285	1.063
<b>Paura</b>	0.235	0.253	,046	1.079	<b>Sorpresa</b>	0.119	0.118	,910	0.992	<b>Rabbia</b>	0.127	0.132	,458	1.041
<b>Rabbia</b>	0.129	0.129	,947	1.004	<b>Paura</b>	0.242	0.240	,851	0.991	<b>Gioia</b>	0.502	0.519	,097	1.035
<b>Tristezza</b>	0.088	0.088	,959	0.996	<b>Gioia</b>	0.512	0.502	,424	0.980	<b>Tristezza</b>	0.087	0.089	,709	1.025
<b>Gioia</b>	0.514	0.502	,297	0.978	<b>Tristezza</b>	0.089	0.086	,694	0.968	<b>Paura</b>	0.248	0.232	,077	0.936
<b>Sorpresa</b>	0.122	0.114	,311	0.941	<b>Rabbia</b>	0.131	0.122	,325	0.936	<b>Disgusto</b>	0.139	0.122	,014	0.874

Livello significatività: 0.05

Livello significatività: 0.10

*Fonte: Elaborazione personale, Excel*

Dall'analisi risulta caratterizzante solamente la “*paura*” per la “*zona rossa*”, conclusione abbastanza coerente e che ci si poteva aspettare, così come il “*disgusto*” è caratterizzante per la non “*zona gialla*”, nel senso che se una persona si trova in “*zona gialla*” è meno probabile che provi “*disgusto*”, anch'esso un risultato coerente.

Inoltre, guardando bene le tabelle, si può notare che il “*disgusto*”, per quanto riguarda la “*zona rossa*”, non è significativo per poco (0,051), è al limite; infatti, se si considerasse un livello di significatività più alto, come  $\alpha=0.10$ , avremmo anche il “*disgusto*” significativo per la “*zona rossa*” e quindi si potrebbe affermare che se un cliente si trova in “*zona rossa*” ha un 11% di probabilità in più di provare “*disgusto*”.

Soffermandoci sempre su un  $\alpha=0.10$ , vediamo come anche “*paura*” e “*gioia*”, per la

“zona gialla”, diventano significative, la prima in modo negativo (chi si trova in “zona gialla” è meno probabile che abbia “paura”) e la seconda in modo positivo (chi si trova in “zona gialla” è più probabile che esprima “gioia”).

In sintesi, potremmo dire che trovarsi in “zona rossa” sembra aumentare le probabilità di avere un cliente con “paura” o “disgusto”, mentre trovarsi in “zona gialla” la probabilità di provare “gioia” con una riduzione della probabilità di avere “paura” o “disgusto”.

Successivamente, l’analisi delle modalità caratterizzanti delle “zone” è stata estesa anche alla variabile “motivazioni” che mostra i seguenti risultati:

*Tabella 18.*

Zona Rossa				Zona Arancione				Zona Gialla				
No	Si	Sign.	Lift	No	Si	Sign.	Lift	No	Si	Sign.	Lift	
Affidabilità sito/app	0.002	0.009	<,001	4.73	Capillarità	0.015	0.026 ,001	1.73	Qualità del servizio	0.068	0.080 ,034	1.17
Accessibilità/Reperibilità	0.094	0.107 ,034		1.15	Qualità del servizio	0.076	0.061 ,027	0.81	Accessibilità/Reperibilità	0.105	0.090 ,023	0.86
Generico	0.247	0.226 ,028		0.92	Affidabilità della banca	0.031	0.022 ,037	0.71	Usability	0.080	0.069 ,049	0.86
					Affidabilità sito/app	0.006	0.001 ,009	0.18	Gamma dei prodotti	0.009	0.006 ,043	0.60
									Affidabilità sito/app	0.006	0.002 ,007	0.39

*Fonte: Elaborazione personale, Excel*

Si può notare subito come in “zona rossa” è presente in modo molto forte “affidabilità sito/app” con una *lift* pari a 4.73 che dice che se una persona si trova in “zona rossa”, la probabilità che citi quel driver aumenta del 373% ([4.73-1] \*100).

Questo, probabilmente, è dovuto al fatto che in “zona rossa” le restrizioni erano tali da non poter uscire di casa, di conseguenza l’unico modo per i clienti di mettersi in

contatto con la banca diventava il mondo dell'online, ossia sito e *app*; stesso discorso per “*accessibilità/reperibilità*”, dato che raramente si poteva andare fisicamente in agenzie.

Le stesse motivazioni si possono trovare analizzando le modalità caratterizzanti della “*zona gialla*”, solo che in questo caso, avendo una misura *lift* <1, indicherebbero che trovarsi in “*zona gialla*” riduce le probabilità di menzionare fattori critici come “*accessibilità/reperibilità*”, “*affidabilità sito/app*” e “*usability*”, mentre aumenta leggermente (circa 17%) la probabilità di parlare della “*qualità del servizio*”. In “*zona arancione*” troviamo grosso modo le stesse tematiche, le quali si comportano in modo simile alla “*zona gialla*”, piuttosto che alle “*zona rossa*”.

Dopo aver visto le modalità caratterizzanti per le “*zone*”, è stata fatta la stessa analisi anche prendendo come riferimento le singole emozioni; di seguito verranno proposti i risultati più significativi per ciascun “*mood*”.

Partendo dalla “*gioia*” si hanno le seguenti modalità caratterizzanti:

*Tabella 69.*

		Gioia	No	Si	Sig.	Lift
<b>DRIVER</b>	<b>Immagine della banca</b>	0.008	0.021	<.001	2.56	
	<b>Qualità del servizio</b>	0.045	0.101	<.001	2.26	
	<b>Affidabilità della banca</b>	0.019	0.038	<.001	2.01	
	<b>ATM</b>	0.038	0.011	<.001	0.28	
	<b>Accessibilità/Reperibilità</b>	0.165	0.035	<.001	0.21	
	<b>Tempistiche</b>	0.123	0.026	<.001	0.21	
<b>PERSONAS</b>	<b>TERESA</b>	0.116	0.144	<.001	1.25	
	<b>NICOLA</b>	0.070	0.086	,005	1.22	
	<b>ANTONIO</b>	0.136	0.108	<.001	0.80	
	<b>VALENTINA</b>	0.078	0.053	<.001	0.67	
<b>LAVORO</b>	<b>DISOCCUPATO</b>	0.014	0.024	<.001	1.70	
	<b>OPERAIO</b>	0.064	0.092	<.001	1.45	
	<b>PROFESSIONISTA AUTONOMO</b>	0.086	0.065	<.001	0.76	
	<b>IMPRENDITORE</b>	0.011	0.007	,044	0.65	

*Fonte: Elaborazione personale, Excel*

Oltre a vedere chi/cosa tende ad incrementare la probabilità di avere la “*gioia*”, è importante vedere anche l’altra parte della medaglia, dove in questo caso abbiamo “*ATM*”, “*accessibilità/reperibilità*” e “*tempistiche*”, drivers che abbassano di molto la probabilità di provare “*gioia*”, segno che qui bisogna fare molto di più per migliorare, poiché nel periodo marzo-dicembre 2020 sono state problematiche emozionali molto forti. Un ottimo lavoro è stato fatto, invece, nella “*qualità dei servizi*”, “*affidabilità della banca*” e “*immagine*” della stessa.

Le persone più semplici sembrerebbe siano state le più felici, si vedano le “*Teresa*” e i “*Nicola*”, accompagnati da “*disoccupati*” e “*operaio*”.

Dopo aver visto la “*gioia*” si è passati all’altro estremo, ossia il “*disgusto*”.

*Tabella 20.*

Disgusto		No	Si	Sig.	Lift
DRIVER	<b>Ambiente filiale</b>	0.015	0.056	<,001	<b>3.71</b>
	<b>Burocrazia</b>	0.016	0.050	<,001	<b>3.12</b>
	<b>ATM</b>	0.020	0.051	<,001	<b>2.56</b>
	<b>Assenza sportello</b>	0.018	0.044	<,001	<b>2.41</b>
	<b>Assistenza chat/telefono</b>	0.037	0.087	<,001	<b>2.35</b>
	<b>Disponibilità sito/app</b>	0.013	0.027	<,001	<b>2.06</b>
	<b>Problem solving/Proattività</b>	0.079	0.053	,002	<b>0.68</b>
	<b>Affidabilità della banca</b>	0.031	0.016	,003	<b>0.51</b>
PERSONAS	<b>VALENTINA</b>	0.060	0.098	<,001	<b>1.62</b>
	<b>NICOLA</b>	0.082	0.055	,001	<b>0.67</b>
	<b>TERESA</b>	0.138	0.077	<,001	<b>0.56</b>
LAVORO	<b>REDDITIERO</b>	0.001	0.006	,001	<b>4.20</b>
	<b>PROFESSIONISTA AUTONOMO</b>	0.073	0.091	,024	<b>1.25</b>
	<b>PENSIONATO</b>	0.356	0.275	<,001	<b>0.77</b>

*Fonte: Elaborazione personale, Excel*

Il “*disgusto*” ha numerose modalità caratterizzanti riguardo i driver (nella tabella sono elencate le più importanti, con lift più elevate), tra tutte “*ambiente filiale*”, “*ATM*”, “*assistenza chat/telefono*” e “*disponibilità sito/app*”; tematiche coerenti con i problemi che il Covid-19 ha portato (rallentamenti *Wi-Fi*, chiusura di attività, etc.).

Coloro che sembrerebbero avere una probabilità maggiore (circa il 320% in più) per essere disgustati sono i “*redditieri*”.

Una situazione simile a quella del “*disgusto*” può essere trovata analizzando la “*rabbia*”; molte tematiche sono in comune:

*Tabella 21.*

<b>Rabbia</b>		No	Si	Sig.	Lift
<b>DRIVER</b>	<b>ATM</b>	0.016	0.081	<,001	<b>5.24</b>
	<b>Disponibilità sito/app</b>	0.010	0.046	<,001	<b>4.42</b>
	<b>Bancomat/Carta di credito</b>	0.012	0.039	<,001	<b>3.10</b>
	<b>Usability</b>	0.062	0.165	<,001	<b>2.66</b>
	<b>Accessibilità/Reperibilità</b>	0.083	0.201	<,001	<b>2.41</b>
	<b>Relazione con DAG</b>	0.007	0.016	,002	<b>2.25</b>
	<b>Immagine della banca</b>	0.017	0.005	,002	<b>0.30</b>
<b>PERSONAS</b>	<b>Affidabilità della banca</b>	0.032	0.008	<,001	<b>0.26</b>
	<b>VALENTINA</b>	0.063	0.083	,008	<b>1.33</b>
	<b>MARCO</b>	0.086	0.104	,046	<b>1.20</b>
<b>LAVORO</b>	<b>NICOLA</b>	0.081	0.058	,005	<b>0.71</b>
	<b>IMPRENDITORE</b>	0.008	0.018	<,001	<b>2.28</b>
	<b>MILITARE</b>	0.070	0.054	,038	<b>0.77</b>
	<b>CASALINGA</b>	0.041	0.027	,019	<b>0.66</b>

*Fonte: Elaborazione personale, Excel*

Possiamo vedere come molte situazioni le troviamo anche nella tabella del “*disgusto*”, una differenza però è l’elevato valore della *lift* di “*ATM*”, che aumenta addirittura del 424% la probabilità di provare “*rabbia*”, seguita da “*disponibilità sito/app*” con 342%. Una distinzione va fatta rispetto al “*disgusto*”, sembrerebbe che la problematica “*ATM*” porti più “*rabbia*” che “*disgusto*”. Cosa significa ciò?

Questo potrebbe essere tradotto sostenendo che il problema “*ATM*” sia un problema temporaneo, poiché la “*rabbia*” può essere vista come un’emozione transitoria (un cliente che trova un *ATM* non funzionante proverebbe “*rabbia*”, ma appena ne trova uno

funzionante quella sensazione di “*rabbia*” potrebbe diminuire), mentre il “*disgusto*” è qualcosa in più, un problema che il cliente potrebbe legarsi al dito e che, di conseguenza, richiede un’azione più forte da parte della banca per poter essere “curato”.

Ciò non toglie che vuol dire che non bisogna agire anche contro la “*rabbia*”.

“*Casalinghe*” e “*militari*” sembrerebbero avere una probabilità minore di provare “*rabbia*” così come i “*Nicola*”.

Successivamente si passa a vedere la situazione della “*paura*”:

*Tabella 22.*

Paura		No	Si	Sig.	Lift
DRIVER	<b>Innovatività</b>	0.003	0.015	<,001	5.76
	<b>Tempistiche</b>	0.040	0.178	<,001	4.49
	<b>Accessibilità/Reperibilità</b>	0.057	0.228	<,001	3.98
	<b>Burocrazia</b>	0.014	0.040	<,001	2.83
	<b>Accesso al credito</b>	0.015	0.036	<,001	2.47
	<b>Rendimento prodotti</b>	0.006	0.013	,001	2.15
PERSONAS	<b>Qualità del servizio</b>	0.087	0.030	<,001	0.35
	<b>VALENTINA</b>	0.061	0.078	,004	1.29
	<b>IVANO</b>	0.161	0.204	<,001	1.27
	<b>FRANCESCO</b>	0.122	0.102	,010	0.83
LAVORO	<b>TERESA</b>	0.138	0.107	<,001	0.78
	<b>PROFESSIONISTA AUTONOMO</b>	0.070	0.092	<,001	1.32
	<b>OPERAIO</b>	0.085	0.056	<,001	0.65

*Fonte: Elaborazione personale, Excel*

Qui sembrerebbe che motivi come “*tempistiche*”, “*innovatività*”, “*accessibilità/reperibilità*” (etc.) aumenterebbero di molto le probabilità di provare “*paura*”. Dando uno sguardo un po’ a tutti i *drivers*, verrebbe da pensare che sono collegati ai soldi, e di conseguenza la “*paura*” potrebbe derivare da una mancanza degli

stessi; si veda “*accesso al credito*”, “*rendimento prodotti*”, ma anche “*tempistiche*” e “*burocrazia*”, dovute magari all’apertura di un mutuo.

Nonostante ciò, ancora una volta le persone più semplici, le quali “*Teresa*” e “*Operai*” tendono ad avere minore probabilità di provare “*paura*”.

Comportamento molto simile alla “*paura*” lo troviamo nella “*tristezza*”:

*Tabella 23.*

		Tristezza	No	Si	Sig.	Lift
<b>DRIVER</b>	<b>Relazione con DAG</b>	0.006	0.031	<,001	<b>5.08</b>	
	<b>Assenza sportello</b>	0.018	0.057	<,001	<b>3.14</b>	
	<b>Rendimento prodotti</b>	0.007	0.015	,011	<b>2.18</b>	
	<b>Burocrazia</b>	0.019	0.038	<,001	<b>2.02</b>	
	<b>Accesso al credito</b>	0.018	0.037	<,001	<b>2.02</b>	
	<b>Cura del cliente</b>	0.148	0.297	<,001	<b>2.01</b>	
<b>PERSONAS</b>	<b>Qualità del servizio</b>	0.076	0.039	<,001	<b>0.51</b>	
	<b>ANTONIO</b>	0.118	0.158	<,001	<b>1.34</b>	
<b>LAVORO</b>	<b>TERESA</b>	0.133	0.097	,003	<b>0.73</b>	
	<b>MANAGER/DIRIGENTE</b>	0.019	0.037	<,001	<b>1.97</b>	
	<b>IMPRENDITORE</b>	0.009	0.016	,042	<b>1.82</b>	
	<b>OPERAIO</b>	0.080	0.055	,011	<b>0.69</b>	
	<b>DISOCCUPATO</b>	0.020	0.006	,005	<b>0.30</b>	

*Fonte: Elaborazione personale, Excel*

Il *driver* che aumenterebbe maggiormente la probabilità di provare “*tristezza*” è “*relazione con DAG*” (408%). In questo caso probabilmente si tratta di persone che vorrebbero essere rassicurate da figure alte della banca, e quindi più competenti; inoltre, anche qui ritroviamo alcune tematiche viste prima come “*rendimento prodotti*”, “*burocrazia*” e “*accesso al credito*”. Ancora una volta i meno tristi sono “*Teresa*”, “*operai*” e anche “*disoccupati*”.

L'ultima situazione da analizzare è quella della “sorpresa”:

*Tabella 24.*

	Sorpresa	No	Si	Sig.	Lift
DRIVER					
	<b>Problem solving/Proattività</b>	0.027	0.432 ,000	<b>15.80</b>	
	<b>Disponibilità/Cortesia</b>	0.104	0.305 <,001	<b>2.94</b>	
	<b>Competenza/Professionalità</b>	0.071	0.160 <,001	<b>2.27</b>	
	<b>ATM</b>	0.026	0.011 ,002	<b>0.42</b>	
	<b>Qualità del servizio</b>	0.080	0.020 <,001	<b>0.25</b>	

*Fonte: Elaborazione personale, Excel*

In questo caso si hanno poche modalità caratterizzanti, dovuto probabilmente alla complessità di questa emozione. Importante però è l'elevato valore di “*problem solving/proattività*”, segno che tale tematica sia risultata un successo; infatti, aumenterebbe del 1480% la probabilità di essere sorpresi. Dall'altra parte della medaglia si hanno “*ATM*” e “*qualità del servizio*”.

Il confronto di tutte le modalità caratterizzanti di ciascuna emozione risulta molto importante, poiché permette di capire effettivamente quali sono state le cause di negatività e quelle di positività.

Dopo aver visto ampliamente e dettagliatamente le modalità caratterizzanti per ciascuna emozione, viene proposta di seguito un’idea di rappresentazione piuttosto “*semplice*” e “*sintetica*”, la quale potrebbe essere molto utile per mostrare i punti di forza<sup>34</sup> di una emozione come la “*gioia*”, o le debolezze di un “*disgusto*”.

Alcune volte questi punti di forza possono essere mostrati direttamente alla clientela per cercare di migliorare ancora di più l’immagine della banca sotto più punti di vista. Mostrare al pubblico, per esempio, che “*affidabilità della banca*” è stata una delle tematiche che ha reso i clienti più gioiosi può in parte ridurre lo scetticismo di quelle persone che vorrebbero avere a che fare con la banca ma, per qualche motivo, non sono sicure proprio sull’affidabilità della stessa.

Lo stesso ragionamento può essere fatto, ovviamente, anche al personale della banca, a cui sarebbe più opportuno mostrare tematiche caratterizzanti emozioni negative, per far capire in modo più visivo ai dipendenti (di ogni grado) le lacune, ossia dove lavorare meglio e migliorare.

---

<sup>34</sup> Nel periodo gennaio 2022 è andato in onda uno spot pubblicitario della *Apple* di forte impatto, che mostrava un episodio in cui una persona cadendo sveniva e il suo *Apple Watch* chiamava i soccorsi inviando la posizione corretta. Un esempio perfetto su come mostrare i punti di forza, attraverso casi reali.

Figura 6. Carta d'identità "gioia"



Fonte: Elaborazione personale, Excel

Figura 7. Carta d'identità "disgusto"



Fonte: Elaborazione personale, Excel

## 2.7 È davvero possibile prevedere un'emozione?

Dopo aver fatto un'ampia analisi descrittiva sul fenomeno di studio, andando a vedere le relazioni tra “*moods*” e altre variabili, quali “*motivazioni*”, “*professioni*”, “*personas*” e altre, si è voluto provare a costruire un modello per cercare di prevedere il “*mood*” avendo a disposizione tutte le altre variabili viste finora come *input*.

Perché l’importanza di un modello di previsione?

L’importanza del modello va cercata nell’analisi fatta fino a questo punto, nelle informazioni che le tabelle in questo elaborato mostravano. Un primo punto fondamentale è sicuramente la relazione vista tra “*moods*” e la scelta di un cliente di rimanere fedele alla banca oppure cercare di meglio altrove; il tipo di professione che un cliente svolge, il segmento di cui fa parte, le motivazioni che menziona, possono essere dei piccoli, ma importanti, segnali che esso stesso ci manda per intervenire, per fare in modo di migliorare la sua situazione, il suo rapporto e le sue problematiche con la banca.

Il vantaggio di sapere, prima che sia troppo tardi, che un cliente ha le caratteristiche tipiche di una persona arrabbiata permette di agire in un determinato modo, al fine di far diventare quella “*rabbia*” in “*gioia*”.

Va ricordato che le emozioni primarie utilizzate in questo elaborato sono sei ed ognuna ha le sue caratteristiche, ciò implica che una persona triste non può essere trattata allo stesso modo di una persona arrabbiata, ogni emozione ha un diverso approccio,

personalizzato e più efficace; inoltre, oltre ad essere tutte diverse tra loro, ciò che le distingue è l'intensità dell'emozione.

Si prenda il caso del “*disgusto*” e della “*rabbia*”, teoricamente verrebbe da pensare che sono emozioni piuttosto simili, trattabili allo stesso modo, e non c’è cosa di più sbagliata. Il “*disgusto*” è qualcosa che va oltre la semplice “*rabbia*”, è una sensazione che si trova in prossimità del punto di non ritorno (si intende del rapporto con la banca), una situazione che deve essere analizzata, “curata” con estrema cautela e dove, soprattutto, non basta una semplice chiamata al cliente ma bisogna fare qualcosa che abbia un grande impatto in modo da far cambiare idea; la “*rabbia*”, invece, può essere vista più come una situazione transitoria, che si può risolvere in una maniera più “*semplice*”.

Per quanto riguarda il modello, come prima cosa sono stato preparato un dataset con tutte le variabili rese dicotomiche (0-1), successivamente è stata provata una regressione multinomiale, avente come output il mood e come input tutte le variabili che sono risultate modalità caratterizzanti nel paragrafo precedente.

Dopo una prima prova, con aggiustamenti riguardo le variabili significative si sono ottenuti i seguenti risultati:

*Tabella 25.*

**Test del rapporto di verosimiglianza**

Effetto	Criteri di adattamento modello Logaritmo della verosimiglianza -2 di modello ridotto	Test del rapporto di verosimiglianza		
		Chi-quadrato	gl	Sign.
Intercetta	10507,808 <sup>a</sup>	,000	0	.
Accessibilità/Reperibilità	10838,161	330,353	5	<,001
Accesso al credito	10529,636	21,828	5	<,001
Affidabilità della banca	10595,088	87,279	5	<,001
Affidabilità/sicurezza sito/app	10521,895	14,087	5	,015
Ambiente filiale	10546,838	39,030	5	<,001
Assenza sportello	10546,754	38,946	5	<,001
ATM	10624,789	116,981	5	<,001
Bancomat/Carta di credito	10542,858	35,050	5	<,001
Burocrazia	10573,241	65,433	5	<,001
Chiarezza informazioni	10521,601	13,792	5	,017
Competenza/Professionalità	10522,465	14,657	5	,012
Costi	10605,929	98,120	5	<,001
Cura del cliente	10545,683	37,874	5	<,001
Disponibilità sito/app	10551,564	43,756	5	<,001
Disponibilità/Cortesia	10638,844	131,036	5	<,001
Generico	11157,235	649,427	5	<,001
Immagine della banca	10566,020	58,212	5	<,001
Innovatività	10534,210	26,402	5	<,001
Modifica condizioni	10534,386	26,577	5	<,001
Problem solving/Proattività	11206,919	699,111	5	<,001
Qualità del servizio	10707,201	199,393	5	<,001
Relazione con DAG	10528,131	20,323	5	,001
Assistenza chat/telefono	10566,540	58,732	5	<,001
Tempistiche	10770,521	262,713	5	<,001
Trasparenza, chiarezza	10519,502	11,694	5	,039
Usability	10582,898	75,090	5	<,001
TERESA	10528,658	20,850	5	<,001
DISOCCUPATO	10529,156	21,348	5	<,001
IMPIEGATO PA	10527,755	19,947	5	,001
MILITARE	10525,998	18,190	5	,003
PENSIONATO	10540,348	32,540	5	<,001
OPERAIO	10532,139	24,330	5	<,001
Zona Gialla	10507,808 <sup>a</sup>	,000	0	.
Zona Arancione	10507,808 <sup>a</sup>	,000	0	.
Zona Rossa	10507,808 <sup>a</sup>	,000	0	.

La statistica chi-quadrato è la differenza nei logaritmi della verosimiglianza -2 tra il modello finale e il modello ridotto. Il modello ridotto viene formato omettendo un effetto dal modello finale. L'ipotesi nulla è che tutti i parametri di quell'effetto siano 0.

a. Questo modello ridotto è equivalente al modello finale perché l'omissione dell'effetto non aumenta i gradi di libertà.

*Fonte: Elaborazione personale, SPSS*

Dalla tabella possiamo vedere come sono risultate significative moltissime variabili;

ciascuna, in parte, contribuisce a migliorare il modello e la previsione. Tra queste abbiamo cinque variabili riguardo la “*professione*”, una variabile che è un cluster della segmentazione “*personas*” e ben ventisei variabili che fanno parte delle “*motivazioni*”.

Il modello, tutto sommato, come prima prova non è andato malissimo e lo possiamo constatare dalle seguenti tabelle.

*Tabella 26.*

<b>Pseudo R-quadrato</b>	
Cox e Snell	,443
Nagelkerke	,465
McFadden	,193

*Fonte: Elaborazione personale, SPSS*

Qui possiamo vedere i valori dello “*Pseudo R-quadrato*”, i quali non vanno interpretati come valori di  $R^2$  (può essere calcolato sono in caso di modelli di regressione lineari)<sup>35</sup>, ma il concetto è lo stesso. I valori stimati da *Cox & Snell*<sup>36</sup> e *Nagelkerke*<sup>37</sup> sono simili, intorno allo 0.4, risultato non altissimo, ma considerando la complessità del tema e considerando che si tratta di una prima prova, tutto sommato poteva andare peggio.

Lo pseudo  $R^2$  di *McFadden*<sup>38</sup> invece, ha un valore letteralmente più basso, 0.193.

A primo impatto potrebbe essere discordante dagli altri due, in realtà è coerente perché empiricamente un valore *McFadden* che oscilla tra 0.2-0.4 indica un ottimo modello; in

<sup>35</sup> L’ $R^2$  indica la bontà di adattamento di un modello, è un valore che oscilla tra 0 e 1

<sup>36</sup>  $R_{C&S}^2 = 1 - (L_0/L_M)^{2/n}$ ,

<sup>37</sup> Lo pseudo  $R^2$  di *Nagelkerke* è una modifica dello pseudo  $R^2$  di *Cox&Snell*, poiché quest’ultimo non può raggiungere valore 1

<sup>38</sup>  $R_{McF}^2 = 1 - \ln(L_M) / \ln(L_0)$ , dove  $L_0$  è il valore della funzione di verosimiglianza del modello senza predittori, mentre  $L_M$  è il valore della verosimiglianza del modello stimato.

questo caso, tende a confermare gli altri due pseudo  $R^2$ .

Bisogna poi vedere la significatività di tale valore, perché uno pseudo  $R^2$  con un alto valore, ma non significativo, non dice assolutamente nulla.

*Tabella 27.*

#### **Informazioni sull'adattamento modello**

Modello	Criteri di adattamento modello Logaritmo della verosimiglianza a -2	Test del rapporto di verosimiglianza		
		Chi-quadrato	gl	Sign.
Solo intercetta	15916,822			
Finale	10507,808	5409,014	170	,000

*Fonte: Elaborazione personale, SPSS*

La tabella, con un valore pari a 0, ci dice che gli pseudo  $R^2$  visti sopra sono significativi.

Successivamente è stato fondamentale andare a vedere l'andamento del modello, ossia come si è comportato nel classificare le varie emozioni; il risultato, anticipato già dallo pseudo  $R^2$  non è dei migliori, ma comunque incoraggiante.

*Tabella 28.*

#### **Classificazione**

Osservato	disgusto	Previsto						Percentuale di correttezza
		gioia	paura	rabbia	sorpresa	tristezza		
disgusto	82	515	298	54	50	3		8,2%
gioia	41	3869	233	52	110	5		89,8%
paura	64	711	735	59	61	4		45,0%
rabbia	47	445	278	110	35	5		12,0%
sorpresa	6	322	107	15	382	0		45,9%
tristezza	20	327	146	31	25	3		0,5%
Percentuale globale	2,8%	66,9%	19,4%	3,5%	7,2%	0,2%		56,0%

*Fonte: Elaborazione personale, SPSS*

La tabella contiene molte informazioni importanti. Sembra che il modello faccia fatica a distinguere e prevedere alcune emozioni, come “*tristezza*” (solo 0.5% corrette), “*disgusto*” (8.2%) e “*rabbia*” (12%). “*Paura*” e “*sorpresa*” hanno una percentuale di correttezza intorno al 45%, in linea ai valori di pseudo  $R^2$  visti. La “*gioia*” sembra essere l’emozione più facilmente interpretabile dal modello, a tratti fin troppo, perché vediamo che tende a classificare molto spesso in “*gioia*”, anche se a volte si tratta di “*moods*” completamente opposti come “*disgusto*”.

Nel complesso la percentuale di classificazione è del 56% e come prima prova, considerando sempre la complessità e l’innovatività del fenomeno, ha dato segnali incoraggianti.

Prima di provare a migliorare il modello, è stata fatta un’analisi più approfondita sul perché il modello tenda a classificare in modo errato, non solo confondendosi tra emozioni negative ma non riuscendo, spesso, a capire nemmeno se si tratti di un’emozione positiva o negativa.

Andando a confrontare i risultati del modello con i “*Verbatim*” dei clienti, effettivamente dei problemi ci sono ed in alcuni casi ha ragione il modello:

Tabella 29.

<b>id</b>	<b>Osservato</b>	<b>Previsto</b>	<b>Classe</b>	<b>Verbatim</b>
95	paura	gioia	Promotore	Ottima app. con la doppia autenticazione e avvisi continui sulle truffe.
2267	rabbia	gioia	Promotore	perche' cliente da circa 40 anni e non posso lamentarmi
3031	rabbia	gioia	Promotore	Talvolta qualcosa non funziona

Fonte: Elaborazione personale, Excel

Questo discorso ci serve per capire che, probabilmente ci sono degli errori già in partenza, ossia nei “*moods*” che avevamo assegnato con l’algoritmo in precedenza, ed in effetti andando a confrontare le emozioni con la *classe NPS* sono venute fuori delle irregolarità piuttosto evidenti:

Tabella 30.

<b>Mood</b>	<b>Detratore</b>	<b>Passivo</b>	<b>Promotore</b>	<b>Totale</b>
<b>disgusto</b>	859	101	42	1002
<b>gioia</b>	681	972	2657	4310
<b>paura</b>	1021	332	281	1634
<b>rabbia</b>	744	111	65	920
<b>sorpresa</b>	287	122	423	832
<b>tristezza</b>	439	81	32	552
<b>Totale complessivo</b>	4031	1719	3500	9250

Fonte: Elaborazione personale, Excel

I valori con lo sfondo colorato non tornano, sono alquanto incoerenti, perché se ci troviamo un cliente “*gioioso*” ci aspettiamo che sia un “*promotore*”, ed invece abbiamo ben 681 persone che sono “*detrattori*”, ma che inizialmente abbiamo classificato con “*gioia*”, ed infatti vedendo alcuni “*verbatim*” notiamo questo:

Tabella 31.

<b>id</b>	<b>Osservato</b>	<b>Previsto</b>	<b>Classe</b>	<b>Verbatim</b>
172	gioia	gioia	Detrattore	non sono soddisfatto
2267	rabbia	gioia	Promotore	perche' cliente da circa 40 anni e non posso lamentarmi
4353	disgusto	gioia	Promotore	Perché sono 36 anni che sono vostro cliente...non credo che devo aggiungere altro.
19261	tristezza	gioia	Promotore	Perché ad oggi la Bnl non mi ha deluso

Fonte: Elaborazione personale, Excel

Negli ultimi tre casi sembra molto evidente la positività dei commenti, inoltre in questo caso il modello aveva previsto bene (*gioia*).

Sembrerebbe poi che ci sia un qualche problema con la parola “*soddisfatto*”, o meglio con le sue negazioni; probabilmente l’algoritmo iniziale leggendo “*soddisfatto*” ha classificato con “*gioia*”, non tenendo conto della negazione iniziale.

Questo discorso sembrerebbe di facile soluzione, magari dicendo all’algoritmo di classificare “*non soddisfatto*” come “*rabbia*” o “*disgusto*”, il problema è che in italiano non si possono considerare tutte le sfaccettature in due casi così banali (“*sono soddisfatto*”, “*non sono soddisfatto*”), perché bisognerebbe analizzare caso per caso dato che le situazioni potrebbero essere tante: “*sono molto soddisfatto*”, “*sono poco soddisfatto*”, “*sono del tutto soddisfatto*”, etc.

Ovviamente questo non vale solo per la parola “*soddisfatto*”, è solamente un esempio per far capire la complessità di ciò che si sta facendo e che in un primo elaborato è abbastanza improbabile riuscire a catturare tutte le possibili sfaccettature; tuttavia, questo vuole essere un punto di partenza, con continui miglioramenti sia da parte dei dati originali (assegnazione del “*mood*” da parte dell’algoritmo) che del modello finale.

Riprendendo il modello, nonostante non funzioni al meglio (classificando il 56% delle emozioni), risulta ugualmente un formidabile classificatore, riuscendo in molti casi addirittura a classificare correttamente l'emozione che in partenza era sbagliata (esempi visti sopra).

Questo significa che quel 56% in realtà potrebbe essere un valore molto più alto, condizionato però dalle emozioni sbagliate già in partenza dal *text miner*, quindi, per provare ad avere una conferma di ciò sono state tolte le osservazioni che nella tabella vista in precedenza risultavano discordanti, insieme ad altre altrettanto contrastanti (“*disgusto*” previsto “*gioia*”, “*rabbia*” previsto “*gioia*” e “*tristezza*” previsto “*gioia*”).

L’ideale sarebbe di andare a vedere manualmente tali casi, uno ad uno (così come quelli visti in precedenza), purtroppo le tempistiche non lo permettono e per semplicità sono stati eliminati, per capire se effettivamente il modello può migliorare andando a ridurre questi elementi di disturbo iniziale

La numerosità campionaria è adesso di 7489 e il modello implementato è il seguente:

*Tabella 32.*

### Pseudo R-quadrato

Cox e Snell	,597
Nagelkerke	,633
McFadden	,317

*Fonte: Elaborazione personale, SPSS*

Lo pseudo  $R^2$  è migliorato notevolmente ed inoltre è ancora significativo.

*Tabella 33.*

Modello	Criteri di adattamento modello Logaritmo della verosimiglianza -2	Test del rapporto di verosimiglianza		
		Chi-quadrato	gl	Sign.
Solo intercetta	14998,417			
Finale	8186,656	6811,761	170	,000

*Fonte: Elaborazione personale, SPSS*

Successivamente andiamo subito a vedere la matrice di confusione:

*Tabella 34.*

Osservato	disgusto	gioia	paura	rabbia	sorpresa	tristezza	Previsto	Percentuale di correttezza
disgusto	92	31	388	40	44	5		15,3%
gioia	18	3278	220	23	89	1		90,3%
paura	85	612	817	52	64	4		50,0%
rabbia	43	26	336	84	32	5		16,0%
sorpresa	11	299	128	13	379	2		45,6%
tristezza	23	9	179	25	26	6		2,2%
Percentuale globale	3,6%	56,8%	27,6%	3,2%	8,5%	0,3%		62,2%

*Fonte: Elaborazione personale, SPSS*

La percentuale adesso è salita fino a 62,2% e potrebbe quasi sicuramente aumentare ancora perché esistono ancora situazioni discordanti; nonostante ciò, si è potuto vedere come migliorando la qualità del dato iniziale, migliori anche il risultato finale. Il modello inizia a capire meglio anche la sostanza delle altre emozioni, con cui faceva più difficoltà precedentemente come “*disgusto*” e “*rabbia*”; la “*tristezza*” resta invece ancora di difficile interpretazione per il modello.

Va ricordato e precisato poi che l'essere umano non è caratterizzato univocamente da un'emozione, nulla vieta di provare più emozioni contemporaneamente, spesso proviamo un mix di emozioni, un misto anche di emozioni positive e negative; infatti, ciò che la regressione multinomiale fa è assegnare una probabilità ad ogni emozione e poi scegliere l'emozione con la probabilità più alta.

Esistono, però, situazioni in cui diventa difficile stabilire l'emozione. Si veda il seguente caso:

*Tabella 35.*

<b>id</b>	<b>Osservato</b>	<b>p_Disgusto</b>	<b>p_Gioia</b>	<b>p_Paura</b>	<b>p_Rabbia</b>	<b>p_Sorpresa</b>	<b>p_Tristezza</b>	<b>Previsto</b>
113	paura	0.242	0.203	0.265	0.100	0.113	0.077	paura

*Fonte: Elaborazione personale, Excel*

Si ha un cliente a cui il modello ha assegnato le seguenti probabilità ad ogni emozione.

Notiamo come non si ha una polarizzazione verso un determinato “*mood*”, anzi, si potrebbe dire che il cliente è composto al 24% di *disgusto*, 20% *gioia*, 27% *paura*, 11% *sorpresa* e 8% *tristezza*. Il modello successivamente assegna “*paura*” perché è dove si ha la probabilità più alta; andando a vedere il commento abbiamo questo:

*Tabella 36.*

<b>id</b>	<b>Verbatim</b>
113	Un'unica pecca tempo eccessivamente lungo nel mettersi in contatto con il vostro call center, sono riuscito dopo 4 tentativi e mezz'ora di attesa.

*Fonte: Elaborazione personale, Excel*

Anche a dando a leggere il commento non risulta facile assegnare un'emozione, si ha un mix tra “*disgusto*” per l'attesa con il call center e “*gioia*” poiché il cliente sottolinea

“un'unica pecca...”, quindi nel complesso è soddisfatto della banca.

Questo per ribadire e mostrare ancora una volta quanto complicata sia questa tematica, il voler associare un'emozione ad una persona.

Per avere una conferma in più sul modello è stata istruita anche una rete neurale sulle 7489 osservazioni e provando varie combinazioni tra “*funzione di attivazione*”<sup>39</sup> e “*n° di strati*”<sup>40</sup> si è avuto il miglior risultato nel seguente caso.

*Tabella 37.*

**Riepilogo elaborazione casi**

	N	Percentuale
Campione	5172	69,1%
Test	2317	30,9%
Valido	7489	100,0%
Escluso	0	
Totali	7489	

*Fonte: Elaborazione personale, SPSS*

Come prima cosa il *dataset* è stato suddiviso in un due campioni, uno per l’addestramento della rete neurale e l’altro per testare il suo funzionamento.

La “*funzione di attivazione*” adoperata è stata una “*sigmoide*”<sup>41</sup>, mentre per *l’output* è stata utilizzata una “*softmax*”<sup>42</sup>.

<sup>39</sup> La funzione di attivazione determina l’output di ciascun neurone

<sup>40</sup> Più strati (o livelli) significano una rete più potente

<sup>41</sup> La funzione *sigmoidea* è una funzione matematica che produce una curva *sigmoide*, ovvero una curva avente un andamento ad "S"

<sup>42</sup> La funzione *softmax* è una generalizzazione di una funzione logistica che comprime un vettore k-dimensionale z di valori reali arbitrari in un vettore k-dimensionale  $\sigma(z)$  di valori compresi in un intervallo (0,1) la cui somma è 1. Viene usata in vari metodi di classificazione multi-classe, come la regressione logistica multinomiale.

I risultati ottenuti rispecchiano molto quelli avuti con la regressione multinomiale con una percentuale di correttezza molto simile.

*Tabella 38.*

<b>Riepilogo del modello</b>		
Addestramento	Errore di entropia incrociata	5249,128
	Percentuale previsioni non corrette	38,5%
	Regola di arresto utilizzata	1 fasi consecutive con nessuna diminuzione dell'errore <sup>a</sup>
Test	Tempo di addestramento	0:00:00,63
	Errore di entropia incrociata	2257,349
	Percentuale previsioni non corrette	37,3%

Variabile dipendente: Mood

a. I calcoli degli errori sono basati sul campione di test.

*Fonte: Elaborazione personale, SPSS*

Per ultimo è stata vista la matrice di confusione, per vedere in cosa ha avuto difficoltà la rete neurale.

*Tabella 39.*

Campione	Osservati	disgusto	Classificazione						Percentuale di correttezza
			gioia	paura	rabbia	sorpresa	tristezza	Previsto	
Addestramento	disgusto	60	48	285	14	37	0	13,5%	
	gioia	14	2227	131	11	64	0	91,0%	
	paura	58	425	606	14	36	0	53,2%	
	rabbia	41	22	247	28	24	0	7,7%	
	sorpresa	14	224	86	4	257	0	43,9%	
	tristezza	15	11	138	8	22	1	0,5%	
	Percentuale globale	3,9%	57,2%	28,9%	1,5%	8,5%	0,0%	61,5%	
Test	disgusto	20	17	102	4	13	0	12,8%	
	gioia	3	1075	77	1	26	0	90,9%	
	paura	23	205	224	13	30	0	45,3%	
	rabbia	17	16	101	15	15	0	9,1%	
	sorpresa	7	79	41	2	118	0	47,8%	
	tristezza	5	2	53	5	8	0	0,0%	
	Percentuale globale	3,2%	60,2%	25,8%	1,7%	9,1%	0,0%	62,7%	

Variabile dipendente: Mood

*Fonte: Elaborazione personale, Excel*

Anche qui ci sono state difficoltà sul tema “tristezza”, seguita da “rabbia” e “disgusto”.

Come visto già prima un problema molto grande risiede direttamente nei dati iniziali, nell'algoritmo usato per assegnare un'emozione, il quale potrà e sarà costantemente migliorato.

Tutto sommato la percentuale di 62.7 raggiunta, considerando che questi sono i primi passi fatti verso questo tipo di studio del cliente, è molto soddisfacente, con la consapevolezza che può notevolmente essere migliorato.

Infine, si è voluto vedere anche un albero di decisione<sup>43</sup>, il quale lavora in modo molto diverso rispetto alla regressione logistica multinomiale e alla rete neurale.

Anche l'albero è stato implementato sul campione di numerosità 7489, ripulito leggermente dalle incoerenze viste sopra; il criterio utilizzato dall'albero è il *CHAID*<sup>44</sup>. Le variabili utilizzate sono le stesse del modello multinomiale della rete, ossia quelle risultate significative dall'analisi delle modalità caratterizzanti.

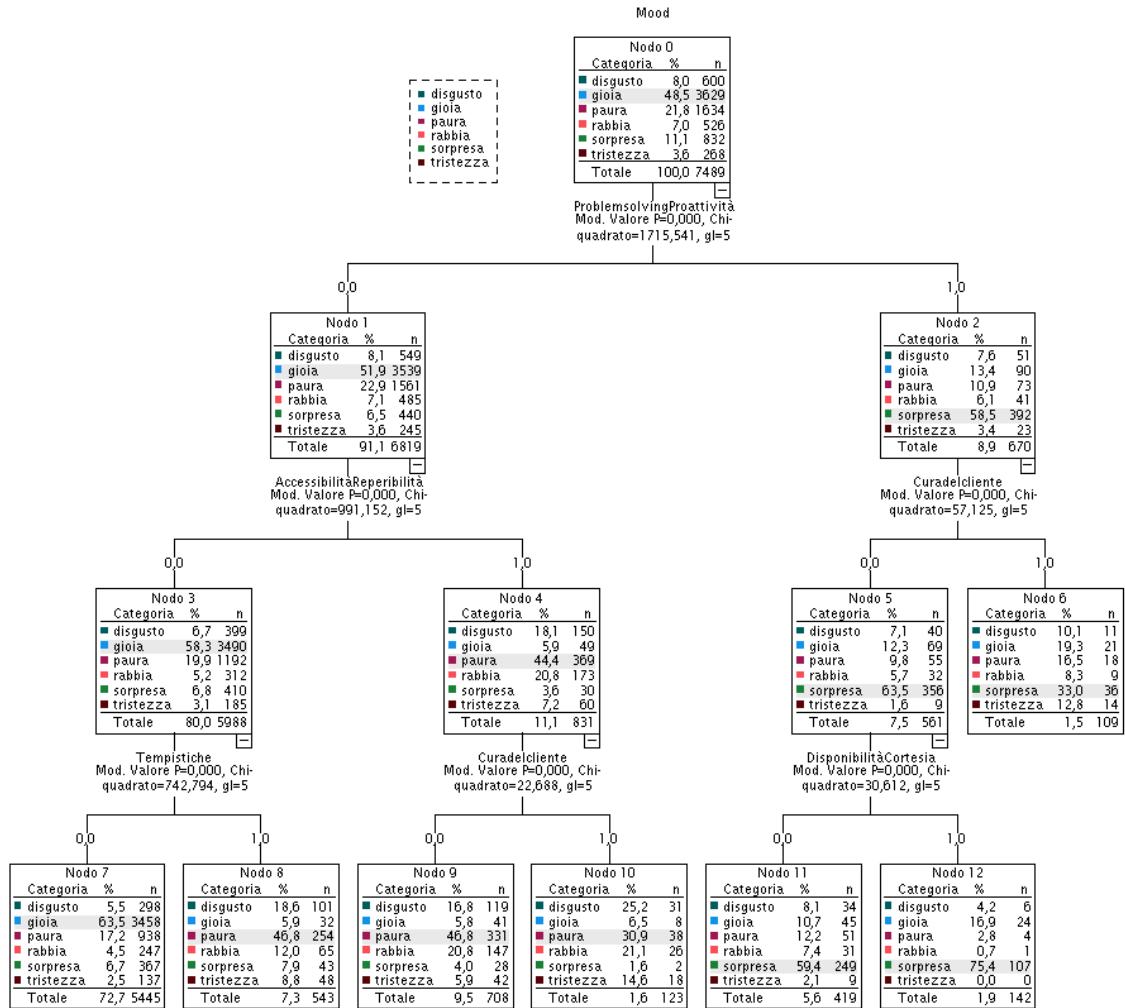
---

<sup>43</sup> Un albero di decisione è un modello predittivo, dove ogni nodo interno (nodo padre) rappresenta una variabile in cui viene fatto un taglio in due nodi figli.

<sup>44</sup> CHAID utilizza un'idea di pre-potatura . Un nodo viene diviso solo se viene soddisfatto un criterio di significatività. Ha lo scopo di lavorare con obiettivi categorici / discretizzati.

Il grafico dell'albero è il seguente:

Figura 8. Albero decisionale



Fonte: Elaborazione personale, SPSS

L'albero utilizza come prima variabile “problem solving/proattività”, mostrando come chi ha menzionato questa motivazione risultò decisamente molto sorpreso con un grande incremento della “sorpresa” e una riduzione abbastanza importante di “gioia” e “paura”. Successivamente, viene considerata dall'albero un'altra motivazione, “cura del cliente”, notando che l'aver menzionato questo driver incide negativamente

aumentando un po' tutte le emozioni negative e con un grande decremento della “*sorpresa*”.

Per coloro che invece non hanno avuto problemi con “*cura del cliente*”, la “*sorpresa*” aumenta ancora di più e le emozioni negative si riducono ancora. Infine, l’albero utilizza come ultimo taglio “*disponibilità/cortesia*” e tra coloro che hanno indicato questo fattore il 75% è sorpreso, segno che è stato fatto un buon lavoro.

Passando al lato sinistro dell’albero, per chi non ha considerato la motivazione “*problem solving/proattività*”, l’albero usa poi come distinzione “*accessibilità/reperibilità*”; qui diventa molto importante il risultato di coloro che hanno nominato il *driver*, poiché ben il 44.4% esprime “*paura*”, con un aumento di tutte le altre emozioni negative e un decremento delle positive. Da qui si è avuto un altro taglio, in base a “*cura del cliente*” che ancora una volta mostra i suoi effetti negativi; infatti, coloro che hanno avuto a che fare con questa tematica la “*paura*” scende da 44.4% a 30.9%, a favore di un incremento molto importante del “*disgusto*” (da 18.1% a 25.2%) e della “*tristezza*” (da 7.2% a 14.6%). Per ultimo, su chi non ha avuto a che fare con “*accessibilità/reperibilità*” è stato fatto un taglio in base alla variabile “*tempistiche*”, confermando anche qui la grande problematica, poiché la “*gioia*” scende da 58.3% a 5.9% a favore del grande incremento di “*paura*” (da 19.9% a 45.8%), seguita dal “*disgusto*” (da 6.7% a 18.6%).

In sintesi, l'albero può essere sintetizzato dicendo che “*cura del cliente*” ha inciso molto negativamente sulla “*tristezza*” e sul “*disgusto*”, “*accessibilità/reperibilità*” su “*disgusto*” e “*paura*” e le “*tempistiche*” molto sulla “*paura*”; per le emozioni positive, “*problem solving/proattività*”, seguita da “*disponibilità/cortesia*” sono stati *drivers* che hanno sorpreso maggiormente i clienti (in positivo).

Per concludere, vediamo come l'albero ha classificato le emozioni:

*Tabella 40.*

Osservato	Previsto						Percentuale di correttezza
	disgusto	gioia	paura	rabbia	sorpresa	tristezza	
disgusto	0	298	251	0	51	0	0,0%
gioia	0	3458	81	0	90	0	95,3%
paura	0	938	623	0	73	0	38,1%
rabbia	0	247	238	0	41	0	0,0%
sorpresa	0	367	73	0	392	0	47,1%
tristezza	0	137	108	0	23	0	0,0%
Percentuale globale	0,0%	72,7%	18,3%	0,0%	8,9%	0,0%	59,7%

Metodo di crescita: CHAID  
Variabile dipendente: Mood

*Fonte: Elaborazione personale, SPSS*

L'albero ha trovato grosse difficoltà con le emozioni “*disgusto*”, “*rabbia*” e “*tristezza*”; infatti, non è riuscito a classificare nemmeno una osservazione ad una di queste tre emozioni. La percentuale di correttezza è del 59,7%, non ottima, ma comunque accettabile per il tema trattato.

Per avere una maggiore sicurezza sui risultati ottenuti dai modelli, sono stati implementati su *Python* sia la regressione multinomiale che l'albero di decisione.

Fortunatamente i risultati ottenuti confermano quelli già in possesso, tranne alcune lievi differenze per quanto riguarda l'albero di decisione, probabilmente perché *Python* di *default* non utilizza la metodologia *CHAID* ma un'altra.

Inoltre, su *Python* è stato implementato anche un algoritmo “*Foresta Casuale*”, che altro non è che un insieme di alberi. Dopo aver ottenuto le previsioni dei diversi alberi sono state calcolate le percentuali di previsioni per ogni albero, per vedere come si è comportato con la classificazione in ciascuno.

I risultati, calcolati poi su *Excel*, sono sintetizzati nella seguente tabella:

*Tabella 41.*

<b>FORESTA CASUALE</b>	
<b>Albero 1</b>	0.50
<b>Albero 2</b>	0.54
<b>Albero 3</b>	0.50
<b>Albero 4</b>	0.53
<b>Albero 5</b>	0.57
<b>Albero 6</b>	0.57
<b>Albero 7</b>	0.53
<b>Albero 8</b>	0.53
<b>Albero 9</b>	0.52
<b>Albero 10</b>	0.53
<b>Media</b>	<b>0.53</b>

*Fonte: Elaborazione personale, Excel*

Andando a vedere invece le previsioni che hanno avuto la prevalenza per ciascun cliente (es. se su un cliente 64, sette alberi hanno classificato “gioia” e tre “sorpresa”, si prende “gioia”) e calcolando la percentuale di classificazione su di esse si ha:

*Tabella 42.*

Oss/Prev	Disgusto	Gioia	Paura	Rabbia	Sorpresa	Tristezza	Totale
<b>Disgusto</b>	0	396	166	0	38	0	<b>0.00</b>
<b>Gioia</b>	0	3492	52	0	85	0	<b>0.96</b>
<b>Paura</b>	0	1197	384	0	53	0	<b>0.24</b>
<b>Rabbia</b>	0	314	185	0	27	0	<b>0.00</b>
<b>Sorpresa</b>	0	427	44	0	361	0	<b>0.43</b>
<b>Tristezza</b>	0	189	62	0	17	0	<b>0.00</b>
<b>Totale</b>	<b>0.00</b>	<b>0.58</b>	<b>0.43</b>	<b>0.00</b>	<b>0.62</b>	<b>0.00</b>	<b>0.57</b>

*Fonte: Elaborazione personale, Excel*

Anche qui non si riesce a prevedere le emozioni “disgusto”, “rabbia” e “tristezza”, come già visto anche per l’albero di classificazione. La percentuale è circa 57%.

Nel caso si volessero consultare personalmente i risultati ottenuti con *Python*, in “Appendice A” sono presenti i codici utilizzati per l’implementazione.

## Conclusioni

L'analisi svolta è iniziata con un semplice calcolo del valore *NPS*, per iniziare a capire in che direzione si stesse muovendo la banca. Una volta fatto ciò si è andato sempre più nello specifico, vedendo l'andamento di tale indice durante l'anno, per capire i periodi più difficili per la banca, e le motivazioni che hanno causato maggiori problemi.

Dopo aver visto le tematiche in cui la banca poteva far di più, attraverso un “*text miner*” sono state estrapolate le emozioni; purtroppo non è stato possibile associare un'emozione ad ogni, ma era un risultato che poteva essere molto prevedibile a causa dell'algoritmo ancora in fase di sviluppo.

Avendo a disposizione le emozioni per una buona parte dei clienti, sono state viste alcune distribuzioni congiunte interessanti, in particolar modo con la segmentazione *personas* e con le professioni. Qui si sono iniziate a vedere le differenze di reazione (in senso emotivo) al periodo *Covid-19* in base proprio alle variabili citate sopra.

La relazione più importante è stata vista in seguito, le emozioni negative tendono a far recensire i clienti in modo molto ostile, aumentando di molto le probabilità di lasciare la banca (o comunque di valutare altre offerte); le emozioni negative il contrario.

Questo risultato ha reso in modo più chiaro il vantaggio, per la banca, di conoscere l'emozione del cliente.

Successivamente è stato creato un indicatore di restrizione (colore della regione), che teneva conto della regione del cliente e della data di intervista, per cercare di

sintetizzare tutte le condizioni esterne (relative al periodo in analisi) in un unico valore; inoltre, per la creazione sono state fatte delle assunzioni poiché il sistema di colorazione delle regioni è entrato in vigore solamente il 6 novembre.

Nonostante ciò, l'indicatore, in un primo momento, ha dato buoni risultati e soprattutto coerenti, se messi in relazione alle emozioni.

Infine, prima di andare a creare un modello di previsione dell'emozione, è stata fatta un'analisi delle modalità caratterizzante per ciascuna emozione, per capire quali fossero le variabili caratterizzanti da usare poi nel modello. I risultati ottenuti qui hanno confermato un po' tutto quello che erano stato visto fino a questo momento.

Per provare a prevedere l'emozione sono stati applicati vari modelli: regressione multinomiale, albero di decisione, rete neurale e foresta casuale.

I risultati, in media, hanno previsto correttamente il 60% delle emozioni, con una previsione massima di 62.7% (rete neurale), un risultato che ha ampliamente soddisfatto le aspettative, non tanto per il valore in sé, ma per ciò che è stato visto durante il percorso di analisi.

È stato visto, infatti, che già in partenza i dati non erano molto attendibili, risultando molto “rumorosi” e di conseguenza andando ad impattare negativamente sulla precisione dei modelli finale. Il problema dei dati rumorosi, però, è tranquillamente risolvibile.

Il *text miner* che è stato utilizzato per assegnare l'emozione è imperfetto, non riesce

ancora a classificare correttamente tutti i casi, sbagliandone anche molti. Nello specifico è stato visto il caso dei “*verbatim*” contenenti la parola “soddisfazione”; qui l’algoritmo non è riuscito a distinguere i commenti in cui si parlava in modo negativo (“non sono soddisfatto”, etc.) da quelli in cui si parlava in modo positivo (“sono molto soddisfatto”, etc.), avendo la tendenza di classificarli in positivo.

Questo difetto potrebbe essere migliorato andando ad inserire nel *text miner* tutti i casi possibili (“sono soddisfatto”, “non sono soddisfatto”, “sono poco soddisfatto”, etc.) per ogni parola inerente al contesto bancario, una procedura che però richiede tempo, dedizione e soprattutto professionalità, poiché un lavoro di questo tipo non può essere fatto da chiunque, ma solamente da esperti psicologi in campo linguistico-emozionale.

Un’altra eventuale soluzione per ottimizzare i risultati è di aggiungere ulteriori informazioni su un cliente, che possono essere ancora più determinanti ed esplicative.

Migliorando la qualità dei dati iniziali i modelli potrebbero avere, quasi sicuramente, una vita meno difficile di quella avuta in questo elaborato, avendo meno rumore.

Ritornando al 62.7% di previsione, è un risultato davvero molto buono e incoraggiante considerando che, oltre alle problematiche già citate, il tema delle emozioni è molto complesso, a volte risulta difficile associare un’emozione anche leggendo il *verbatim*.

Questo elaborato, quindi, non vuole essere un *finish*, ma uno *start* per questo tema così complicato delle emozioni che, in un futuro non troppo lontano, potrebbe diventare meno intricato.

# Appendice A: Codici Python

## Modelli

### Regressione multinomiale

```
In [1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn import tree
from sklearn.tree import DecisionTreeClassifier
from sklearn.linear_model import LogisticRegression
import seaborn as sns
```

```
In [2]: #carico il file
data = pd.read_excel("C:\\\\Users\\\\emanu\\\\Desktop\\\\Multinomiale_finale.xlsx", sheet_name="Data_Regressi
print(data.shape)

(7489, 37)
```

```
In [3]: import warnings
warnings.filterwarnings("ignore")
```

```
In [4]: data.columns
```

```
Out[4]: Index(['id', 'Mood', 'AccessibilitàReperibilità', 'Accessoalcredito',
   'Affidabilitàdellabanca', 'Affidabilitàsicurezzasitoapp',
   'Ambientefiliale', 'Assenzasportello', 'ATM', 'BancomatCartadicredito',
   'Burocrazia', 'Chiarezzainformazioni', 'CompetenzaProfessionalità',
   'Costi', 'Curadelcliente', 'Disponibilitàsitoapp',
   'DisponibilitàCortesia', 'Generico', 'Immaginedellabanca',
   'Innovatività', 'Modificacondizioni', 'ProblemsolvingProattività',
   'Qualitàdelservizio', 'RelazioneconDAG', 'Assistenzachattelefono',
   'Tempistiche', 'Trasparenzachiarezza', 'Usability', 'TERESA',
   'IMPIEGATOPA', 'IMPRENDITORE', 'MILITARE', 'OPERAIO', 'PENSIONATO',
   'ZonaGialla', 'ZonaArancione', 'ZonaRossa'],
  dtype='object')
```

```
In [5]: y = data.loc[:, 'Mood']
x = data.loc[:, 'AccessibilitàReperibilità':]
```

```
In [6]: data.loc[:, 'Mood'].value_counts()
```

```
Out[6]: gioia      3629
paura      1634
sorpresa     832
disgusto     600
rabbia      526
tristezza    268
Name: Mood, dtype: int64
```

```
In [7]: Nomi_variabili = data.columns[2:]
Nomi_variabili
```

```
Out[7]: Index(['AccessibilitàReperibilità', 'Accessoalcredito',
   'Affidabilitàdellabanca', 'Affidabilitàsicurezzasitoapp',
   'Ambientefiliale', 'Assenzasportello', 'ATM', 'BancomatCartadicredito',
   'Burocrazia', 'Chiarezzainformazioni', 'CompetenzaProfessionalità',
   'Costi', 'Curadelcliente', 'Disponibilitàsitoapp',
   'DisponibilitàCortesia', 'Generico', 'Immaginedellabanca',
   'Innovatività', 'Modificacondizioni', 'ProblemsolvingProattività',
   'Qualitàdelservizio', 'RelazioneconDAG', 'Assistenzachattelefono',
   'Tempistiche', 'Trasparenzachiarezza', 'Usability', 'TERESA',
   'IMPIEGATOPA', 'IMPRENDITORE', 'MILITARE', 'OPERAIO', 'PENSIONATO',
   'ZonaGialla', 'ZonaArancione', 'ZonaRossa'],
  dtype='object')
```

```
In [8]: lr = LogisticRegression(multi_class='multinomial')
```

```
In [9]: param_range = [100, 10, 1, 0.1, 0.001, 0.0000001, 0.000000000000001]

from sklearn.model_selection import validation_curve
train_scores, test_scores = validation_curve(estimator=lr, X=x, y=y, param_name='C',
                                             param_range=param_range, cv=10)

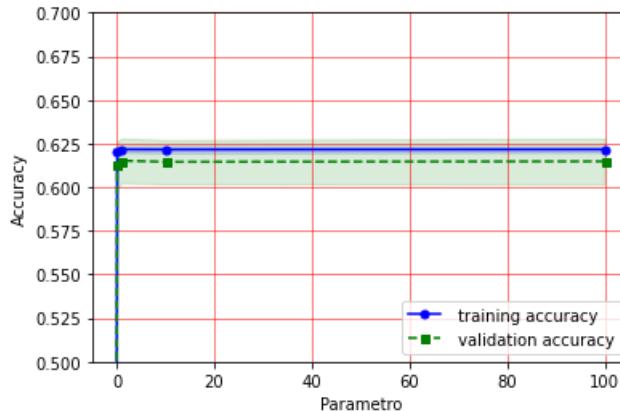
train_mean = np.mean(train_scores, axis=1)
train_std = np.std(train_scores, axis=1)
test_std = np.std(test_scores, axis=1)
test_mean = np.mean(test_scores, axis=1)

plt.plot(param_range, train_mean, color="blue", marker="o", markersize=5, label="training accuracy")
plt.fill_between(param_range, train_mean + train_std, train_mean - train_std, alpha=0.15, color="blue")

plt.plot(param_range, test_mean, color="green", linestyle="--", marker="s", markersize=5, label="validation accuracy")
plt.fill_between(param_range, test_mean + test_std, test_mean - test_std, alpha=0.15, color="green")

plt.grid(color='r', linewidth=0.5)
plt.legend(loc="lower right")
plt.xlabel("Parametro")
plt.ylabel("Accuracy")

plt.ylim([0.5, 0.7])
plt.show()
```



```
In [10]: lr.fit(x,y)
lr.score(x,y)
```

```
Out[10]: 0.6218453732140473
```

```
In [11]: lr.predict_proba(x)
```

```
Out[11]: array([[1.45182008e-03, 9.29932691e-01, 5.00917666e-02, 9.13881810e-04,
   1.67465773e-02, 8.63263485e-04],
   [1.06808919e-02, 8.76514648e-01, 8.38615293e-02, 2.85363823e-03,
   2.27578591e-02, 3.33143365e-03],
   [2.26843548e-01, 6.10631683e-03, 2.21968123e-01, 3.60281782e-01,
   5.52911399e-03, 1.79271116e-01],
   ...,
   [4.06469013e-01, 8.40723368e-03, 3.17538877e-01, 1.46933573e-01,
   1.59770620e-02, 1.04674241e-01],
   [5.70116862e-02, 4.66871265e-01, 2.94941308e-01, 5.00494286e-02,
   8.98127851e-02, 4.13135268e-02],
   [4.27342440e-01, 4.55777950e-03, 9.37404799e-02, 1.71130002e-01,
   2.14799625e-01, 8.84296732e-02]])
```

```
In [12]: data['previsione_disgusto'] = lr.predict_proba(x)[:,0]
data['previsione_gioia'] = lr.predict_proba(x)[:,1]
data['previsione_paura'] = lr.predict_proba(x)[:,2]
data['previsione_rabbia'] = lr.predict_proba(x)[:,3]
data['previsione_sorpresa'] = lr.predict_proba(x)[:,4]
data['previsione_tristezza'] = lr.predict_proba(x)[:,5]
data['previsione'] = lr.predict(x)
data.head()
```

```
Out[12]: comatCartadicredito ... ZonaGialla ZonaArancione ZonaRossa previsione_disgusto previsione_gioia previsione_paura previs
      0 ... 0 0 1 0.001452 0.929933 0.050092
      0 ... 0 0 1 0.010681 0.876515 0.083862
      0 ... 0 0 1 0.226844 0.006106 0.221968
      0 ... 0 0 1 0.018307 0.864024 0.086931
      0 ... 0 0 1 0.001452 0.929933 0.050092
```

In [13]: lr.coef\_

```
Out[13]: array([[ 8.08100982e-01,  6.90692241e-01, -2.73547225e-01,
   2.05552628e-01,  1.20038547e+00,  7.81049816e-01,
   8.16973032e-01,  7.65133224e-01,  1.12900237e+00,
   3.65305601e-01, -1.98588535e-01,  1.06019212e+00,
  -4.92418481e-02,  3.03008838e-01, -4.51789529e-01,
  -1.57555998e+00, -2.29159143e-01, -5.09937704e-01,
   5.05343806e-01,  1.67957106e-01, -1.24519958e-01,
   5.70827478e-01,  8.17915825e-01,  8.47970097e-01,
   2.38010067e-01, -2.60783920e-02, -8.72425431e-01,
  -7.20987381e-01, -1.78963659e-01, -6.47642058e-01,
  -1.34865929e-01, -3.30000561e-01, -1.31226496e-01,
   1.48002592e-02,  1.08840558e-01],
 [-2.12933543e+00, -1.34053655e+00,  1.16152156e+00,
   7.69261819e-01, -7.51310761e-01, -1.24140154e+00,
  -1.67620630e+00, -1.89586119e+00, -1.97811366e+00,
  -3.56764460e-01,  1.78745803e-01, -5.27639984e-01,
  -5.12228908e-02, -7.10714672e-01,  4.87359409e-01,
   3.06861705e+00,  1.58792841e+00, -4.52512015e-01,
  -1.43616935e+00, -1.18347714e+00,  1.91165697e+00,
  -1.36511919e+00, -1.36906234e+00, -2.26998389e+00,
  -1.23701988e+00,  2.40344412e-01,  3.40283067e-01,
   1.44873159e-01,  5.20419103e-02,  2.15423844e-01,
   4.76496482e-01,  2.23202397e-01,  1.13725665e-02,
   4.93839755e-02, -3.52215255e-02],
 [ 3.63511539e-01,  6.99328902e-02,  2.75405589e-01,
   2.32402456e-01,  8.63326322e-02, -1.65120712e-01,
  -5.24694126e-01, -2.00106200e-01,  1.00754548e-01,
  -1.78182357e-01, -1.23407671e-01,  1.59046363e-02,
  -1.28392805e-03, -6.00064306e-01, -1.62815445e-01,
   5.38130872e-01,  1.79368741e-01,  8.18548053e-01,
   5.43850799e-02, -8.56556427e-01,  5.94507180e-03,
  -1.03955915e+00, -2.41940557e-01,  4.69616106e-01,
  -8.22605644e-02,  2.57668968e-01, -6.00891841e-02,
  -1.58381351e-01, -1.84426000e-01,  6.22241491e-02,
  -1.35135254e-01,  1.72897230e-01, -3.60332837e-02,
   1.74476362e-02,  3.58193890e-02],
 [ 1.28692755e+00,  3.94783604e-01, -9.53299455e-01,
   1.56450241e-01,  2.93390264e-01,  5.93569984e-01,
   1.64546138e+00,  1.32331322e+00,  8.58692343e-01,
   1.18537488e-01, -3.54654460e-03,  2.71354230e-01,
  -4.73736552e-02,  1.42945143e+00, -6.14994160e-01,
  -1.48467291e+00, -1.19460140e+00,  2.87318409e-01,
   1.18058061e+00,  7.35769601e-02, -1.08092489e+00,
   8.05018931e-01,  3.99557227e-01,  6.20214422e-01,
   5.40208033e-01,  4.95179357e-01,  1.41762337e-01,
   1.16194922e-01,  5.93271154e-01, -1.49036830e-02,
   1.86380571e-01, -1.39698132e-01,  1.37650847e-01,
  -1.07817851e-01, -4.39278178e-02],
 [-9.74392091e-01, -6.52248234e-01, -2.31563430e-01,
  -8.10585157e-01, -8.31833830e-01, -1.13542130e+00,
  -1.16018716e+00, -5.73134653e-01, -9.58502449e-01,
  -6.21790157e-01,  3.45185904e-01, -1.33488450e+00,
  -3.04277489e-01, -6.68294711e-01,  1.01503862e+00,
   5.88376114e-01, -8.84388536e-02, -5.46514106e-01,
  -1.37831498e+00,  1.71035337e+00, -2.32120383e-01,
  -4.29750143e-01, -3.43623715e-02, -5.39144944e-01,
  -6.43238076e-01, -3.30479748e-01,  3.08638818e-01,
   5.83332589e-02, -1.95975872e-01,  4.70256920e-01,
   1.38140794e-01,  2.52614245e-01,  7.60344494e-02,
   3.84706343e-02, -1.11965393e-01],
 [ 6.45187456e-01,  8.37376045e-01,  2.14829616e-02,
  -5.53081987e-01,  3.03622055e-03,  1.16732375e+00,
   8.98653177e-01,  5.80655598e-01,  8.48166845e-01,
   6.72893886e-01, -1.98388956e-01,  5.15073494e-01,
   4.53399811e-01,  2.46613418e-01, -2.72798899e-01,
  -1.13489116e+00, -2.55097753e-01,  4.03097363e-01,
   1.07417482e+00,  8.81461251e-02, -4.80036813e-01,
   1.45858208e+00,  4.27892212e-01,  8.71328215e-01,
   1.18430042e+00, -6.36634598e-01,  1.41830394e-01,
   5.59967393e-01, -8.59475326e-02, -8.53591718e-02,
  -5.31016663e-01, -1.79015179e-01, -5.77980833e-02,
  -1.22846538e-02,  4.64547894e-02]])
```

```
In [14]: Odds = np.exp(lr.coef_)
```

```
In [15]: data_odds = pd.DataFrame()
data_odds['Variabili'] = Nomi_variabili
data_odds['Odds'] = Odds[0]
data_odds
```

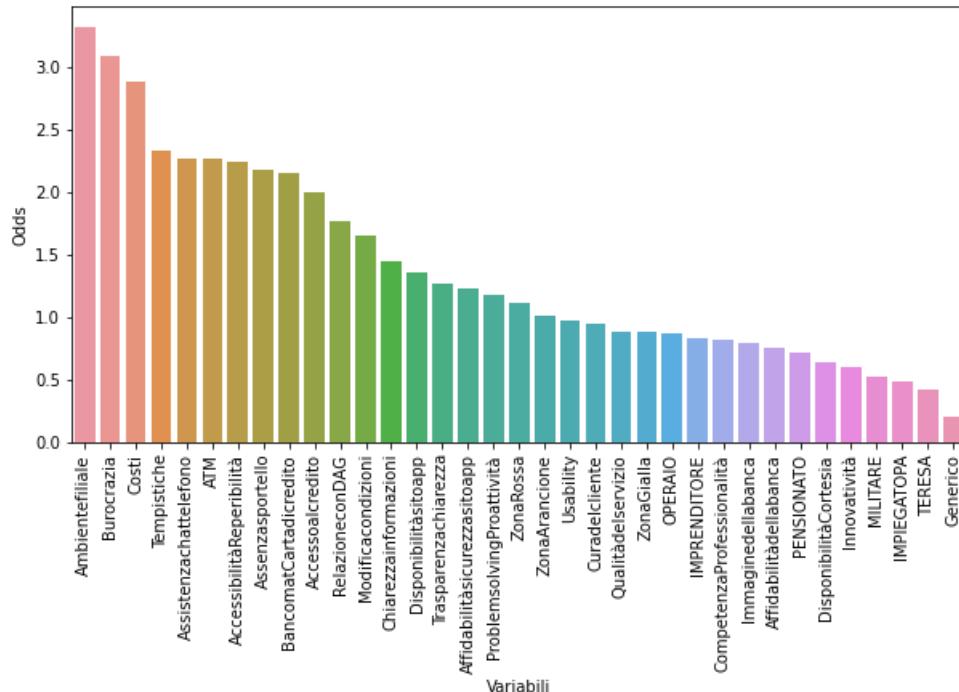
Out[15]:

	Variabili	Odds
0	AccessibilitàReperibilità	2.243643
1	Accessoalcredito	1.995096
2	Affidabilitàdellabanca	0.760676
3	Affidabilitàsicurezzasitoapp	1.228204
4	Ambientefiliale	3.321397
5	Assenzasportello	2.183764
6	ATM	2.263637
7	BancomatCartadicredito	2.149281
8	Burocrazia	3.092570
9	Chiarezzainformazioni	1.440954
10	CompetenzaProfessionalità	0.819887
11	Costi	2.886926
12	Curadelcliente	0.951951
13	Disponibilitàsitoapp	1.353926
14	DisponibilitàCortesia	0.636488
15	Generico	0.206892
16	Immaginedellabanca	0.795202
17	Innovatività	0.600533
18	Modificacondizioni	1.657555
19	ProblemsolvingProattività	1.182886
20	Qualitàdelservizio	0.882921
21	RelazioneconDAG	1.769731
22	Assistenzachattelefono	2.265773
23	Tempistiche	2.334902
24	Trasparenzachiarezza	1.268722
25	Usability	0.974259
26	TERESA	0.417937
27	IMPIEGATOPA	0.486272
28	IMPRENDITORE	0.836136
29	MILITARE	0.523278
30	OPERAIO	0.873833
31	PENSIONATO	0.718923
32	ZonaGialla	0.877019
33	ZonaArancione	1.014910
34	ZonaRossa	1.114985

```
In [16]: data_odds.sort_values(by = ['Odds'], inplace=True, ascending = False)
plt.figure(figsize = (10,5))

plt.xticks (rotation = 90)
sns. barplot(x = 'Variabili', y = 'Odds', data = data_odds)
```

Out[16]: <AxesSubplot:xlabel='Variabili', ylabel='Odds'>



```
In [17]: from sklearn.metrics import confusion_matrix

matrix = confusion_matrix(data.Mood,data.previsione)
def plot_confusion_matrix(cm, classes, normalize=False, title='Confusion matrix', cmap=plt.cm.Blues)
    """
    This function prints and plots the confusion matrix.
    Normalization can be applied by setting `normalize=True`.
    """

    if normalize:
        cm = cm.astype('float') / cm.sum(axis=1)[:, np.newaxis]

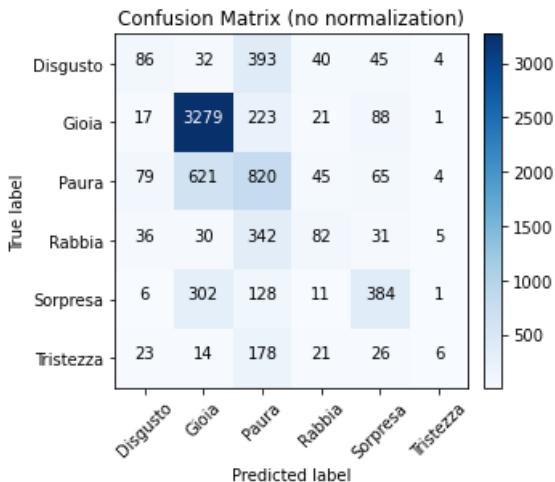
    im = plt.imshow(cm, interpolation='nearest', cmap=cmap)
    plt.title(title)
    plt.colorbar(im, fraction=0.046, pad=0.04)
    tick_marks = np.arange(len(classes))
    plt.xticks(tick_marks, classes, rotation=45)
    plt.yticks(tick_marks, classes)

    fmt = '.2f' if normalize else 'd'
    thresh = cm.max() / 2.
    for i, j in itertools.product(range(cm.shape[0]), range(cm.shape[1])):
        plt.text(j, i, format(cm[i, j], fmt),
                 horizontalalignment="center",
                 color="white" if cm[i, j] > thresh else "black")

    plt.ylabel('True label')
    plt.xlabel('Predicted label')
    plt.tight_layout()
```

```
In [18]: import itertools

plt.figure(figsize=(5,5))
plot_confusion_matrix(matrix, classes=['Disgusto', 'Gioia', 'Paura', 'Rabbia', 'Sorpresa', 'Tristezza'],
                      title='Confusion Matrix (no normalization)')
plt.show()
```



```
In [19]: from sklearn.metrics import classification_report
print(classification_report(data.Mood, data.previsione))
```

	precision	recall	f1-score	support
disgusto	0.35	0.14	0.20	600
gioia	0.77	0.90	0.83	3629
paura	0.39	0.50	0.44	1634
rabbia	0.37	0.16	0.22	526
sorpresa	0.60	0.46	0.52	832
tristezza	0.29	0.02	0.04	268
accuracy			0.62	7489
macro avg	0.46	0.36	0.38	7489
weighted avg	0.59	0.62	0.59	7489

## Albero decisionale

```
In [20]: #carico il file
data = pd.read_excel("C:\\\\Users\\\\emanu\\\\Desktop\\\\Multinomiale_finale.xlsx", sheet_name="Data_Albero")
print(data.shape)
```

(7489, 36)

```
In [21]: data.columns
```

```
Out[21]: Index(['id', 'Mood', 'AccessibilitàReperibilità', 'Accessoalcredito',
   'Affidabilitàdellabanca', 'Affidabilitàsicurezzasitoapp',
   'Ambienteefiliale', 'Assenzasportello', 'ATM', 'BancomatCartadicredito',
   'Burocrazia', 'Chiarezzainformazioni', 'CompetenzaProfessionalità',
   'Costi', 'Curadelcliente', 'Disponibilitàsitoapp',
   'DisponibilitàCortesia', 'Immaginedellabanca', 'Innovatività',
   'Modificacondizioni', 'ProblemsolvingProattività', 'Qualitàdelservizio',
   'RelazioneconDAG', 'Assistenzachattelefono', 'Tempistiche',
   'Trasparenzachiarezza', 'Usability', 'TERESA', 'IMPIEGATOPA',
   'IMPRENDITORE', 'MILITARE', 'OPERAIO', 'PENSIONATO', 'ZonaGialla',
   'ZonaArancione', 'ZonaRossa'],
  dtype='object')
```

```
In [22]: y = data.loc[:, 'Mood']
x = data.loc[:, 'AccessibilitàReperibilità':]
```

```
In [23]: Nomi_variabili = data.columns[2:]
Nomi_variabili
```

```
Out[23]: Index(['AccessibilitàReperibilità', 'Accessoalcredito',
   'Affidabilitàdellabanca', 'Affidabilitàsicurezzasitoapp',
   'Ambienteefiliale', 'Assenzasportello', 'ATM', 'BancomatCartadicredito',
   'Burocrazia', 'Chiarezzainformazioni', 'CompetenzaProfessionalità',
   'Costi', 'Curadelcliente', 'Disponibilitàsitoapp',
   'DisponibilitàCortesia', 'Immaginedellabanca', 'Innovatività',
   'Modificacondizioni', 'ProblemsolvingProattività', 'Qualitàdelservizio',
   'RelazioneconDAG', 'Assistenzachattelefono', 'Tempistiche',
   'Trasparenzachiarezza', 'Usability', 'TERESA', 'IMPIEGATOPA',
   'IMPRENDITORE', 'MILITARE', 'OPERAIO', 'PENSIONATO', 'ZonaGialla',
   'ZonaArancione', 'ZonaRossa'],
  dtype='object')
```

```
In [24]: albero = DecisionTreeClassifier(max_depth= 3)
albero.fit(x,y)
albero.score(x,y)
```

Out[24]: 0.5978101215115502

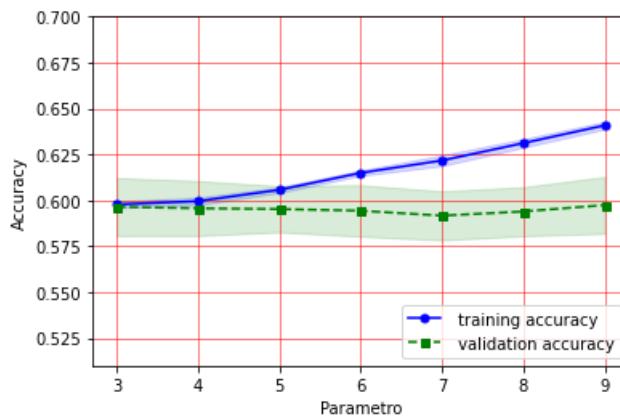
```
In [25]: param_range = [3,4,5,6,7,8,9]
from sklearn.model_selection import validation_curve
train_scores, test_scores = validation_curve( estimator= albero, X = x, y = y,param_name='max_depth',
                                              param_range=param_range, cv=10)

train_mean = np.mean( train_scores, axis=1)
train_std=np.std(train_scores, axis=1)
test_std= np.std(test_scores, axis=1)
test_mean= np.mean(test_scores, axis=1)

plt.plot(param_range, train_mean, color="blue", marker="o", markersize=5, label=" training accuracy")
plt.fill_between(param_range, train_mean + train_std, train_mean - train_std, alpha=0.15, color="blue")

plt.plot(param_range, test_mean, color="green", linestyle="--",marker="s", markersize=5, label=" vali")
plt.fill_between(param_range, test_mean + test_std, test_mean - test_std, alpha=0.15, color="green")

plt.grid(color='r', linewidth=0.5)
plt.legend( loc="lower right")
plt.xlabel ("Parametro")
plt.ylabel ("Accuracy")
plt.ylim([0.51,0.7])
plt.show()
```



```
In [26]: importanza = albero.feature_importances_
importanza
```

```
Out[26]: array([0.34939611, 0.          , 0.          , 0.          , 0.          ,
       0.          , 0.          , 0.00271354, 0.          , 0.          ,
       0.          , 0.00748217, 0.01354834, 0.          , 0.          ,
       0.          , 0.          , 0.          , 0.33804556, 0.          ,
       0.          , 0.          , 0.28162875, 0.          , 0.          ,
       0.          , 0.          , 0.          , 0.          , 0.          ,
       0.00718552, 0.          , 0.          , 0.          ])
```

```
In [27]: Importanza_variabili= pd.DataFrame()
Importanza_variabili ['Variabili'] = Nomi_variabili
Importanza_variabili ['Importanza'] = importanza
```

```
In [28]: data['Previsione_Albero_disgusto'] = albero.predict_proba(x)[:,0]
data['Previsione_Albero_gioia'] = albero.predict_proba(x)[:,1]
data['Previsione_Albero_paura'] = albero.predict_proba(x)[:,2]
data['Previsione_Albero_rabbia'] = albero.predict_proba(x)[:,3]
data['Previsione_Albero_sorpresa'] = albero.predict_proba(x)[:,4]
data['Previsione_Albero_tristezza'] = albero.predict_proba(x)[:,5]
data['Previsione_Albero'] = albero.predict(x)
```

In [29]: Importanza\_variabili

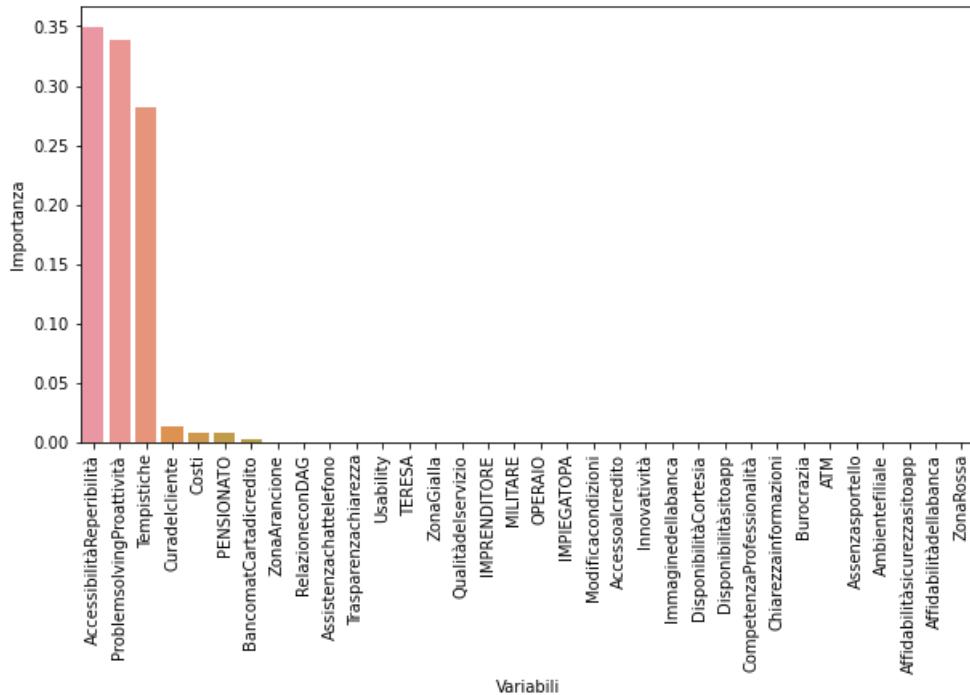
Out[29]:

	Variabili	Importanza
0	AccessibilitàReperibilità	0.349396
1	Accessoalcredito	0.000000
2	Affidabilitàdellabanca	0.000000
3	Affidabilitàsicurezzasitoapp	0.000000
4	Ambientefiliale	0.000000
5	Assenzasportello	0.000000
6	ATM	0.000000
7	BancomatCartadicredito	0.002714
8	Burocrazia	0.000000
9	Chiarezzainformazioni	0.000000
10	CompetenzaProfessionalità	0.000000
11	Costi	0.007482
12	Curadelcliente	0.013548
13	Disponibilitàsitoapp	0.000000
14	DisponibilitàCortesia	0.000000
15	Immaginedellabanca	0.000000
16	Innovatività	0.000000
17	Modificacondizioni	0.000000
18	ProblemsolvingProattività	0.338046
19	Qualitàdelservizio	0.000000
20	RelazioneconDAG	0.000000
21	Assistenzachattelefono	0.000000
22	Tempistiche	0.281629
23	Trasparenzachiarezza	0.000000
24	Usability	0.000000
25	TERESA	0.000000
26	IMPIEGATOPA	0.000000
27	IMPRENDITORE	0.000000
28	MILITARE	0.000000
29	OPERAIO	0.000000
30	PENSIONATO	0.007186
31	ZonaGialla	0.000000
32	ZonaArancione	0.000000
33	ZonaRossa	0.000000

```
In [30]: Importanza_variabili.sort_values(by = ['Importanza'], inplace=True, ascending = False)
plt.figure(figsize = (10,5))

plt.xticks(rotation = 90)
sns.barplot(x = 'Variabili', y = 'Importanza', data = Importanza_variabili)

Out[30]: <AxesSubplot:xlabel='Variabili', ylabel='Importanza'>
```



```
In [31]: from sklearn.metrics import confusion_matrix

matrix = confusion_matrix( data.Mood,data.Previsione_Albero)
def plot_confusion_matrix(cm, classes, normalize=False, title='Confusion matrix', cmap=plt.cm.Blues)
    """
    This function prints and plots the confusion matrix.
    Normalization can be applied by setting `normalize=True`.
    """

    if normalize:
        cm = cm.astype('float') / cm.sum(axis=1)[:, np.newaxis]

    im = plt.imshow(cm, interpolation='nearest', cmap=cmap)
    plt.title(title)
    plt.colorbar(im, fraction=0.046, pad=0.04)
    tick_marks = np.arange(len(classes))
    plt.xticks(tick_marks, classes, rotation=45)
    plt.yticks(tick_marks, classes)

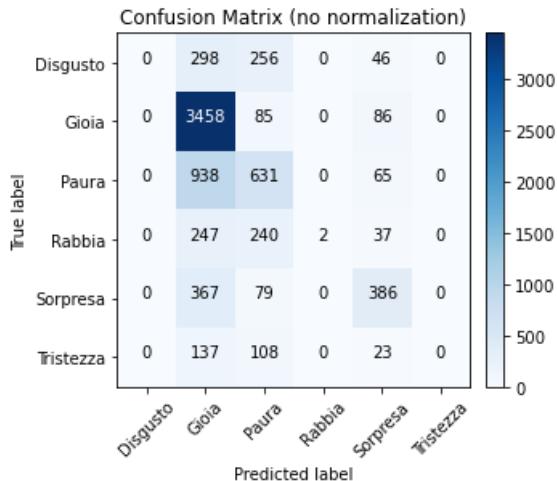
    fmt = '.2f' if normalize else 'd'
    thresh = cm.max() / 2.
    for i, j in itertools.product(range(cm.shape[0]), range(cm.shape[1])):
        plt.text(j, i, format(cm[i, j], fmt),
                 horizontalalignment="center",
                 color="white" if cm[i, j] > thresh else "black")

    plt.ylabel('True label')
    plt.xlabel('Predicted label')
    plt.tight_layout()
```

```
In [32]: import itertools

plt.figure(figsize=(5,5))
plot_confusion_matrix(matrix, classes=['Disgusto','Gioia', 'Paura', 'Rabbia', 'Sorpresa', 'Tristezza'],
                      title='Confusion Matrix (no normalization)')

plt.show()
```



```
In [33]: from sklearn.metrics import classification_report
print(classification_report(data.Mood, data.Previsione_Albero))
```

	precision	recall	f1-score	support
disgusto	0.00	0.00	0.00	600
gioia	0.64	0.95	0.76	3629
paura	0.45	0.39	0.42	1634
rabbia	1.00	0.00	0.01	526
sorpresa	0.60	0.46	0.52	832
tristezza	0.00	0.00	0.00	268
accuracy			0.60	7489
macro avg	0.45	0.30	0.28	7489
weighted avg	0.54	0.60	0.52	7489

## Visualizzare Albero

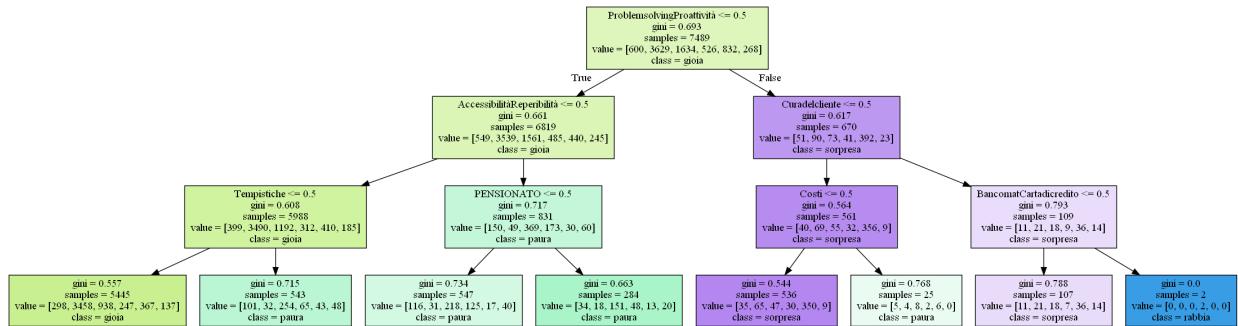
```
In [34]: from sklearn.tree import export_graphviz
from IPython.display import SVG
import graphviz
from IPython.display import display
```

```
In [35]: from sklearn import tree
import pydotplus

dot_data = tree.export_graphviz(albero, out_file=None,
                               feature_names=Nomi_variabili, class_names=['disgusto', 'gioia', 'paura', 'rabbia', 'sorpresa'],
                               filled = True)
graph = pydotplus.graph_from_dot_data(dot_data)
```

```
In [36]: from IPython.display import Image
graph = pydotplus.graph_from_dot_data(dot_data)
Image(graph.create_png())
```

Out[36]:



## Foreste casuali

```
In [37]: from sklearn.ensemble import RandomForestClassifier
```

```
In [38]: #carico il file
data = pd.read_excel("C:\\\\Users\\\\emanu\\\\Desktop\\\\Multinomiale_finale.xlsx", sheet_name="Data_Albero")
print(data.shape)

(7489, 36)
```

```
In [39]: y = data.loc[:, 'Mood']
x = data.loc[:, 'AccessibilitàReperibilità':]
```

```
In [40]: Nomi_variabili = data.columns[2:]
Nomi_variabili
```

```
Out[40]: Index(['AccessibilitàReperibilità', 'Accessoalcredito',
               'Affidabilitàdellabanca', 'Affidabilitàsicurezzasitoapp',
               'Ambientefiliale', 'Assenzasportello', 'ATM', 'BancomatCartadicredito',
               'Burocrazia', 'Chiarezzainformazioni', 'CompetenzaProfessionalità',
               'Costi', 'Curadelcliente', 'Disponibilitàsitoapp',
               'DisponibilitàCortesia', 'Immaginedellabanca', 'Innovatività',
               'Modificacondizioni', 'ProblemsolvingProattività', 'Qualitàdelservizio',
               'RelazioneconDAG', 'Assistenzachattelefono', 'Tempistiche',
               'Trasparenzachiarezza', 'Usability', 'TERESA', 'IMPIEGATOPA',
               'IMPRENDITORE', 'MILITARE', 'OPERAIO', 'PENSIONATO', 'ZonaGialla',
               'ZonaArancione', 'ZonaRossa'],
              dtype='object')
```

```
In [41]: foresta = RandomForestClassifier(n_estimators = 10, max_depth= 3)
foresta.fit(x,y)
foresta.score(x,y)
```

Out[41]: 0.5310455334490586

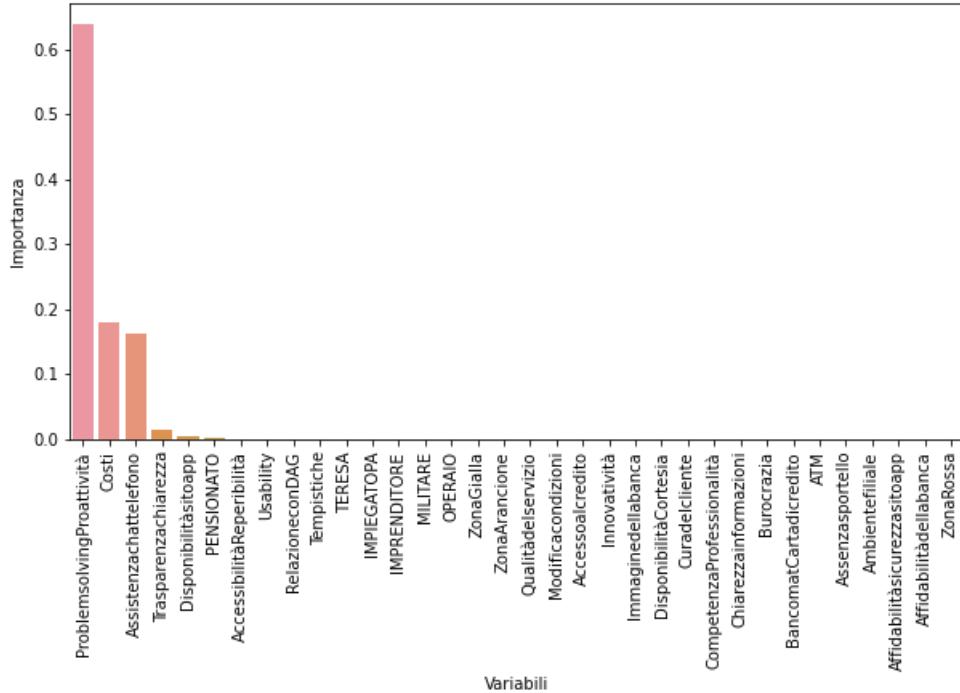
```
In [42]: Importanza = foresta.estimators_[0].feature_importances_
```

```
In [43]: Importanza_variabili= pd.DataFrame()
Importanza_variabili ['Variabili'] = Nomi_variabili
Importanza_variabili ['Importanza'] = Importanza
```

```
In [44]: Importanza_variabili.sort_values(by = ['Importanza'], inplace=True, ascending = False)
plt.figure(figsize = (10,5))

plt.xticks (rotation = 90)
sns.barplot(x = 'Variabili', y = 'Importanza', data = Importanza_variabili)
```

```
Out[44]: <AxesSubplot:xlabel='Variabili', ylabel='Importanza'>
```



```
In [45]: c=0
c=c+1
for tree_in_forest in foresta.estimators_:
    c+=1
    variabile = "albero_forest_" + str(c)
    data[variabile] = tree_in_forest.predict(x)
```

```
In [46]: data.head()
```

```
Out[46]: tadicredito ... albero_forest_2 albero_forest_3 albero_forest_4 albero_forest_5 albero_forest_6 albero_forest_7 albero_forest_8
0 ... 1.0 1.0 1.0 1.0 1.0 1.0 1.0
0 ... 1.0 1.0 1.0 1.0 1.0 1.0 1.0
0 ... 2.0 2.0 2.0 2.0 2.0 1.0 1.0
0 ... 1.0 1.0 1.0 1.0 1.0 1.0 1.0
0 ... 1.0 1.0 1.0 1.0 1.0 1.0 1.0
```

```
In [14]: data.to_csv('foresta.csv')
```

## Text Miner

```
In [3]: import numpy as np
import pandas as pd
import nltk
from sklearn import feature_extraction
from nltk.stem.snowball import SnowballStemmer
import re
stemmer = SnowballStemmer("italian")
pd.set_option('display.max_columns',200)
pd.options.display.float_format = '{:,.0f}'.format
```

```
In [4]: #carico il file
data = pd.read_excel("C:\\\\Users\\\\Family\\\\Desktop\\\\Ema\\\\Python\\\\Dataset.xlsx")
print(data.shape)

(45713, 31)
```

```
In [5]: data = data[["NPS(TM)", "Verbatim"]]
data.head()
```

	NPS(TM)	Verbatim
0	8	facilità d'uso
1	8	Procedere con app o home banking è abbastanza ...
2	9	Tutto
3	8	Perché sino più di 40 anni che la mia famiglia...
4	4	Dipendenti non sempre gentili. Competitor più...

```
In [6]: def NPScluster(x):
    if (x<7):
        x = 'Detrattore'
    elif ((x>6) & (x<9)):
        x = 'Neutrale'
    else:
        x = 'Promotore'
    return x

def tokenize_and_stem(text):
    text = text.lower()
    tokens = [word for sent in nltk.sent_tokenize(text) for word in nltk.word_tokenize(sent)]
    filtered_tokens = []
    for token in tokens:
        if re.search('[a-zA-Z]', token):
            filtered_tokens.append(token)
    stems = [stemmer.stem(t) for t in filtered_tokens]
    return stems

#metodo di matching
def matchCompost(verbatim,parole):
    import re
    result = False
    i = 0
    while(i<len(parole)):
        if(re.search(parole[i],str(verbatim), re.IGNORECASE)):
            result = True
            i+=1
        else:
            i+=1
    return result

#1
def creaColonneDriverVerbatim(data,driver):
    lenDriver = len(driver)
    for i in range(lenDriver):
        data = riempিColonneDriverVerbatim(data,driver[i])
    return data

#2
def riempিColonneDriverVerbatim(data, driver):
```

```

data[driver] = data['Verbatim'].apply(lambda x: driver01(str(x),driver))
return data

#3
def driver01(verbatim, driver):
    dizionario = corrispondenzeDriver(verbatim,driver)
    if dizionario.get(driver) == driver:
        return 1
    else:
        return 0
#4
def corrispondenzeDriver(verbatim,driver):
    d0="" "#rabbia"
    d1="" "#disgusto"
    d2="" "#tristezza"
    d3="" "#gioia"
    d4="" "#paura"
    d5="" "#sorpresa"

    cont=0
    verbatim = verbatim.lower()

    l0=[" dissens"," non approv"," conflitt"," non d'accord", " inaccord"," divergenz"," lacinant"," mal di
        " nessuna risposta", " non risponde nessuno", " nessuno rispond", " frustrat", " non riesco", " frustrant
        " non si può", " non mi soddisf", " insoddisfatt", " inappagat", " afflitt", " macello", " casino", " lament
        " lamentel", " non funzion", " blocco", " arrabbiat", " mi arrabbiar", " incazza", " rott", " stanc", " ira"
        " al limite", " non gentilezz", " complessit", " peggi", " peggior", " poco professional", "assoluta manc
        " disappunto", " non possibil", " non soddi", " mai trov ben", "non ti aiut", " irrispett", " insoddisfat

    l1=[" mult", " sanzion", " scorrett", " inutil", " dev", " indignat", " pazzi", " assurd", " non tutelat", " schif
        " disapprov", " ma che cazz", " che cavol", " disprezz", " vergogn", " pessim", " eccessiv", " ridicol", " s
        " scadent"]

    l2=[" amareggiat", " delus", " fastid", " sfiduci", " preoccupat", " senza fiducia", " non ho fiducia", " contra
        " quando riprenderà", " spento", " trist", " malinconic", " pensieros", " passat", " megli prim", " inquiet
        " non so cos far", " annoiat", " noia", " deprim", " deprimment", " stess cos", " ripetitiv", " routin", " nos
        " giù corda", " infelic", " desperat", " tristezz", " piangere", " abbattut", " amarezz", " abandon", " sco
        " dispiac", " mancanz", " manca"]

    l3=[" vicin", " sentiment", " affezionat", " benevolenz", " tenerezz", " affett", " attaccament", " affettuosit
        " mi siete vicini", " vicin", " sostegn", " prezios", " protegge", " amiciz", " amic", " friend", " ammirazio
        " tutel", " creder", " fidat", " fidar", " mi fido", " fiduc", " riconoscenz", " grato", " grata", " gratitudi
        " trovo ben", " soddisfaz", " soddisfatt", " stimat", " considerat", " d'accord", " rispett", " tutel", " ma
        " mai lament", " non mi mai delus", " non mi delud", " aiut"]

    l4=[" drammatic", " fa tremar", " moment difficol", " moment difficil", " non facil", " emergenz", " non sicur
        " depress", " groppo in gola", " ansi", " oppressio", " oppress", " panico", " desperat", " timor", " tem",
        " truff", " difficol", " difficil", " incub", " non vicin", " nonostant", " preoccup", ]

    l5=[" entusiast", " eccitazion", " appagat", " fogat", " sorpres", " oltre le aspettativ", " accoglienz", " stupi
        " prontezz", " risolve", " risolt", " soluzion", " scopert", " agi"]

    if(driver=="rabbia"):
        if matchCompost(verbatim, l0):
            d0="rabbia"
            cont = +1
    if(driver=="disgusto"):
        if matchCompost(verbatim, l1):
            d1="disgusto"
            cont = +1
    if(driver=="tristezza"):
        if matchCompost(verbatim, l2):
            d2="tristezza"
            cont = +1
    if(driver=="gioia"):
        if matchCompost(verbatim, l3):
            d3="gioia"
            cont = +1
    if(driver=="paura"):
        if matchCompost(verbatim, l4):
            d4="paura"
            cont = +1
    if(driver=="sorpresa"):
        if matchCompost(verbatim, l5):
            d5="sorpresa"
            cont = +1

    dictDriver = {"rabbia":d0,
                  "disgusto":d1,
                  "tristezza":d2,

```

```

        "gioia":d3,
        "paura":d4,
        "sorpresa":d5}
# print(cont)
return dictDriver

```

In [8]:

```

drivers = ["rabbia",
           "disgusto",
           "tristezza",
           "gioia",
           "paura",
           "sorpresa"]

```

In [10]:

```

%%time
#data = data[data.Verbatim.str.contains("app")]
dataWork = creaColonneDriverVerbatim(data,drivers)
dataWork['generico'] = dataWork.loc[:, "rabbia":].sum(axis=1)
dataWork['generico'] = dataWork.generico.apply(lambda x: 1 if x==0 else 0)
dataWork.head()

```

Wall time: 23 s

Out[10]:

	NPS(TM)	Verbatim	rabbia	disgusto	tristezza	gioia	paura	sorpresa	generico
0	8	facilità d'uso	0	0	0	0	0	0	1
1	8	Procedere con app o home banking è abbastanza ...	0	0	0	0	0	0	1
2	9	Tutto	0	0	0	0	0	0	1
3	8	Perché sino più di 40 anni che la mia famiglia...	0	0	0	1	0	0	0
4	4	Dipendenti non sempre gentili. Competitor più...	0	0	0	0	0	0	1

## Bibliografia

- AUSTRALIAN HUMAN RIGHTS COMMISSION (2 August 1995). «*Bringing them Home - preliminary*».
- BAEHRE, S., O'DWYER, M., O'MALLEY, L., & LEE, N. (2021). «*The use of Net Promoter Score (NPS) to predict sales growth: insights from an empirical investigation*», Journal of the Academy of Marketing Science.
- BORZACCHIELLO P. (agosto 2021). «*Basta dirlo-Le parole da scegliere e le parole da evitare per una vita felice*», Mondadori Libri S.p.A., Milano, I edizione.
- COWEN, A. S., & KELTNER, D. (2017). «*Self-report captures 27 distinct categories of emotion bridged by continuous gradients*», PNAS, September 5, 2017.
- EKMAN, P. (1992). «*Are There Basic Emotions?*». Psychological Review, 99(3), 550-553.
- EL NAQA, I., & MURPHY, M. J. «*What is Machine Learning?*», pp 3-11.
- GAZZEI, D. S., & BORRELLI, F. (2016). «*Net Promoter Score: solo soddisfazione e raccomandabilità?*», MK n°6.
- LEWIS, C., & MEHMET, M. (2020). «*Does the NPS® reflect consumer sentiment? A qualitative examination of the NPS using a sentiment analysis approach*», Internatiol Journal of Market Research.

- MARTIN, J. R., & WHITE, P. R. R. (2004). «*The language of evaluation: Appraisal in English*», Berlin, Germany: Springer, «*Strategy Maps*», Harvard Business School Press.
- ONDAATJE, M. (2000). «*Anil's Ghost*», Toronto, Vintage.
- PELLEGRINO G., FADIGA L., FOGASSI L., GALLESE V., & RIZZOLATTI G. (1992). «*Understanding motor events: a neurophysiological study*». *Exp. Brain Res.* 91: 176-180.
- READ, J., & CARROLL, J. (2010). «*Annotating expressions of Appraisal in English*», Springerlink.
- REICHHELD, F. F. (2003). «*The one number you need to grow*». *Harvard Business Review*, 81, 46–55.
- RIZZOLATTI, G., & SINIGAGLIA, C. (2006). «*So quel che fai. Il cervello che agisce e i neuroni specchio*», Milano, Raffaello Cortina Editore.

## Sitografia

- A list apart, «*Dall'empatia all'advocacy*»,  
Url: <https://alistapart.com/it/article/dall-empatia-all-advocacy/>
- IlSole24Ore, «*La pandemia che ha sconvolte le nostre vite e resterà per sempre nell'immaginario comune. Una cronistoria degli eventi che non avremmo mai potuto immaginare*», Url: <https://lab24.ilsole24ore.com/storia-coronavirus/>
- Inside Marketing, «*Impatto delle emozioni sulla decisione di acquisto*», Url: <https://www.insidemarketing.it/neuromarketing-emozioni-guida-decisioni-di-acquisto/>
- Marketing Journal, «*Marketing e comunicazione empatica: saper ascoltare per saper emozionare*», Url: <https://www.marketingjournal.it/marketing-e-comunicazione-empatica-saper-ascoltare-per-saper-emozionare/>
- Psychology Today, «*Empathy*»,  
Url: <https://www.psychologytoday.com/us/basics/empathy>
- QAStack, Url: <https://qastack.it/>
- State of Mind (Il giornale delle scienze psicologiche), «*Empatia*»,  
Url: <https://www.stateofmind.it/tag/empatia/>

- StatisticalHorizons, «*What's the Best R-Squared for Logistic Regression?*», Url: <https://statisticalhorizons.com/r2logistic>
- Wikipedia, «*Advocacy*»,  
Url: <https://it.wikipedia.org/wiki/Advocacy>
- Wikipedia, «*Neuromarketing*»,  
Url: <https://it.wikipedia.org/wiki/Neuromarketing>
- YouTrend, «*Un anno di zone colorate*»,  
Url: <https://www.youtrend.it/2021/11/10/un-anno-di-zone-colorate/>