



Insurance Data Science Marketing Actuary

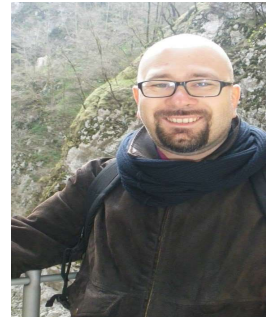
Claudio Giorgio Giancaterino, Intesa SanPaolo Vita



About the speaker

- **Claudio G. Giancaterino** – *Actuary, Intesa SanPaolo Vita*

Actuary and data science enthusiast. I have always looked towards the future with a keen interest in innovation. This has led me to work with numbers as an actuary. In my free time, in recent years, I have devoted time to the field of data science. I started with competitions on Kaggle and other hackathons, then I increased my involvement by teaching practical lessons on insurance statistics at the Catholic University of Milan, writing articles for Data Science Milan meetup, and holding workshops or presentations at conferences.



Intesa Sanpaolo Vita, is the Insurance Group of the Intesa Sanpaolo banking group.

It was born in 2012 from the merger of EurizonVita, Intesa Vita, Sud Polo Vita and Centrovita Assicurazioni. In 2021 merged Assicurazioni Vita, Lombarda Vita and BAP.

Intesa Sanpaolo Vita offers insurance products and pension savings solutions that create value for the customer thanks to a constant care at the service level with an increasing attention along the various phases of life of the contract.





Introduction

Actuaries have traditionally been associated with the insurance industry, where they are used to estimate risks, insurance premiums, and reserving. However, in recent years, actuarial skills have found application in a number of other fields, including marketing



**“Insurance Companies need to leverage
data science and marketing
to improve
actuarial assumptions
and optimize
business performance“**



“Integrating data science and marketing in actuarial assumptions can improve business performance and customer experience”

Customer Segmentation

- **Actuaries can use unsupervised learning to group customers based on their features, preferences and behaviours identifying patterns and trends by the Customer Portfolio analysis**
- **Benefits: better understanding of customer needs, preferences and potential, tailored marketing strategies, increased customer loyalty and retention**



Data Analysis

- **Actuarial skills can be used for:**
 - ❖ **Customer data analysis**
 - ❖ **Market trend analysis**
 - ❖ **Sales data analysis**
- **Benefits: data-driven decision-making, improved targeting, enhanced profitability, and more effective resource allocation, leading to better marketing outcomes**



Pricing Strategies

- **Actuaries can use their expert judgment in the process of combining data, models and experience to make informed decisions. They can develop pricing models and services that reflect customer value, market trends and risks**
- **Benefits: optimized pricing strategies, increased profitability, competitive advantage**



Predictive Models

- **Actuaries can use supervised learning to predict events or indicators that affect marketing performance, such as churn or customer lifetime value**
- **Benefits: improved forecasting accuracy, enhanced marketing effectiveness, reduced costs and waste**



Risk Assessment

- **Actuaries can use their expert judgment assessing risks in marketing campaigns, analyzing customer behaviour risks, evaluating market trend risks, allocating marketing resources effectively, achieving the best return on investment**
- **Benefits: help insurers make more informed business decisions, improve operational efficiency, and increase profitability**



Case Studies:

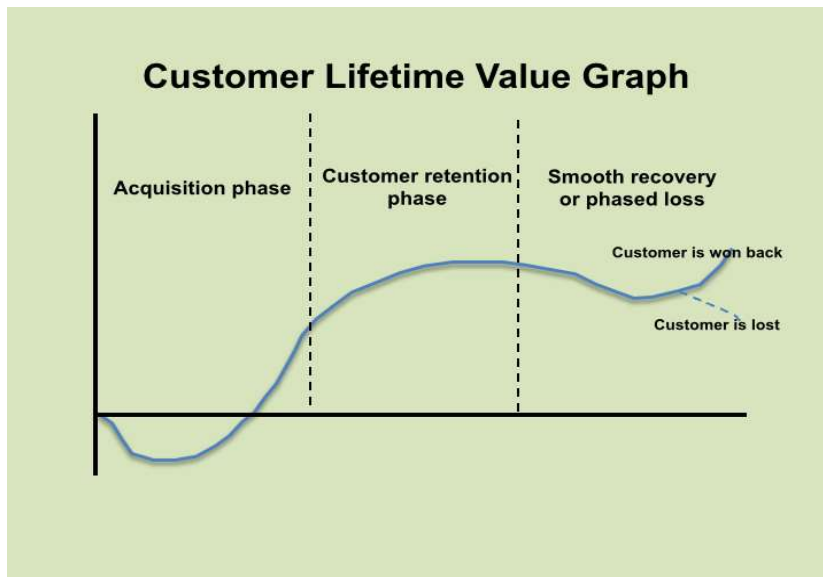
- **Customer Lifetime Value Prediction and Segmentation**
- **Cross-Selling Assessment**
- **Customer Sentiment Analysis**

Customer Lifetime Value Prediction & Segmentation

Go to the article



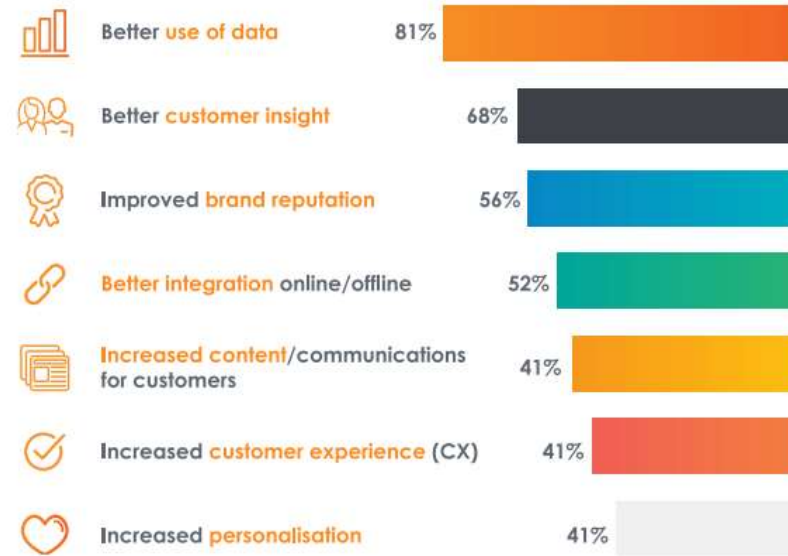
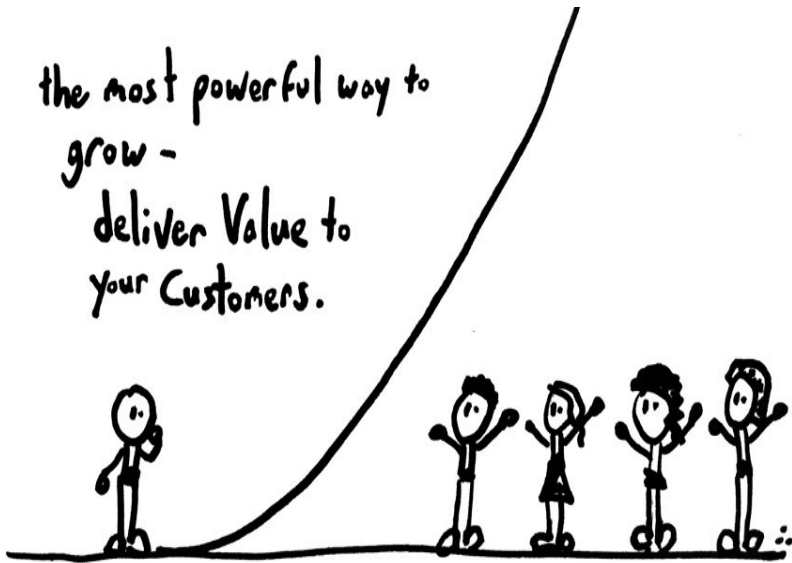
What is Customer Lifetime Value?
It's the value of a customer for a Company over the span of their lifetime relationship.



businessgrowthguys.com

$$CLV_i = \sum_{t=1}^N \frac{(p_{i,t} - c_{i,t})pr_{i,t}}{(1 + dr)^t} - AC_i$$

the most powerful way to grow -
 deliver Value to
 your Customers.



What increases CLV in business, according to a survey by Criteo ([source](#))

Use case: AutoInsurance Data Set

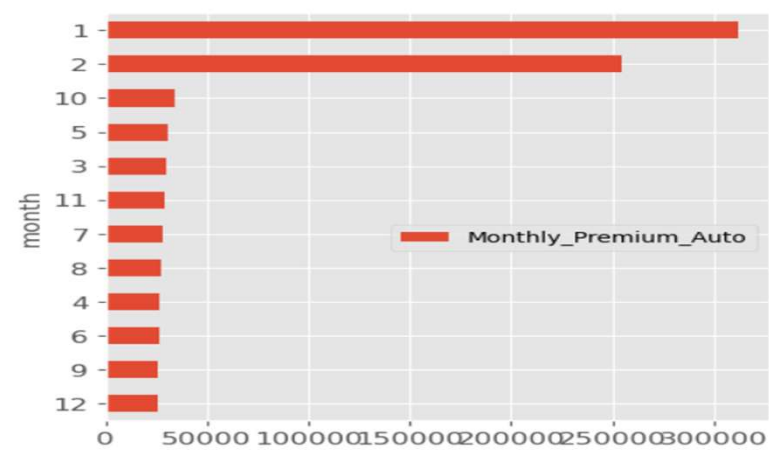
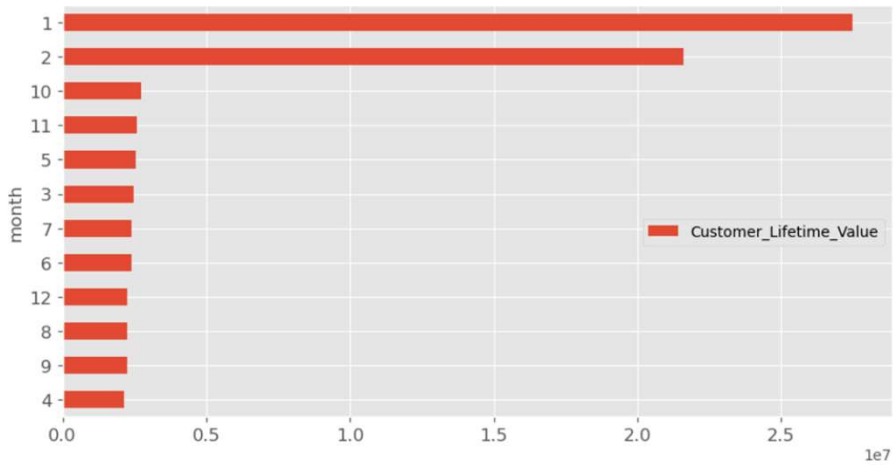
	count	unique	top	freq
State	9134	5	California	3150
Response	9134	2	No	7826
Coverage	9134	3	Basic	5568
Education	9134	5	Bachelor	2748
EmploymentStatus	9134	5	Employed	5698
Gender	9134	2	F	4658
Location_Code	9134	3	Suburban	5779
Marital_Status	9134	3	Married	5298
Policy_Type	9134	3	Personal Auto	6788
Policy	9134	9	Personal L3	3426
Renew_Offer_Type	9134	4	Offer1	3752
Sales_Channel	9134	4	Agent	3477
Vehicle_Class	9134	6	Four-Door Car	4621
Vehicle_Size	9134	3	Medsize	6424

	count	mean	std	min	25%	50%	75%	max
Customer_Lifetime_Value	9134.0	8004.940475	6870.967608	1898.007675	3994.251794	5780.182197	8962.167041	83325.381190
Income	9134.0	37657.380009	30379.904734	0.000000	0.000000	33889.500000	62320.000000	99981.000000
Monthly_Premium_Auto	9134.0	93.219291	34.407967	61.000000	68.000000	83.000000	109.000000	298.000000
Months_Since_Last_Claim	9134.0	15.097000	10.073257	0.000000	6.000000	14.000000	23.000000	35.000000
Months_Since_Policy_Inception	9134.0	48.064594	27.905991	0.000000	24.000000	48.000000	71.000000	99.000000
Number_of_Open_Complaints	9134.0	0.384388	0.910384	0.000000	0.000000	0.000000	0.000000	5.000000
Number_of_Policies	9134.0	2.966170	2.390182	1.000000	1.000000	2.000000	4.000000	9.000000
Total_Claim_Amount	9134.0	434.088794	290.500092	0.099007	272.258244	383.945434	547.514839	2893.239678
eoymissing_days	9134.0	276.667068	93.820412	29.000000	244.000000	315.000000	340.000000	364.000000

Distribution of CLV and Premium Auto in the year

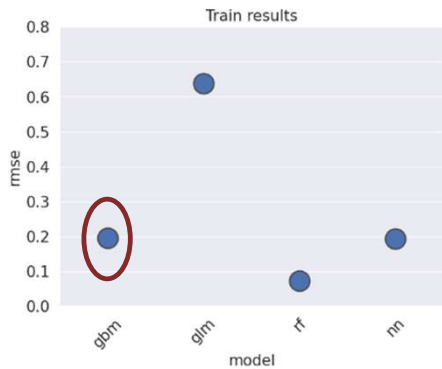
	month	Customer_Lifetime_Value	PERCENTAGE
0	1	27,465,144.97	37.563217
1	2	21,607,247.52	29.551555
2	10	2,737,806.65	3.744412
3	11	2,601,649.84	3.558195
4	5	2,546,155.61	3.482297
5	3	2,460,476.93	3.365117
6	7	2,398,846.49	3.280827
7	6	2,386,698.89	3.264213
8	12	2,268,604.26	3.102699
9	8	2,255,473.06	3.084740
10	9	2,246,222.23	3.072088
11	4	2,142,799.86	2.930640

	month	Monthly_Premium_Auto	PERCENTAGE
0	1	311,719.00	36.609726
1	2	254,472.00	29.886372
2	10	34,505.00	4.052427
3	5	31,167.00	3.660397
4	3	29,814.00	3.501494
5	11	28,941.00	3.398965
6	7	28,218.00	3.314053
7	8	27,201.00	3.194612
8	4	26,964.00	3.166777
9	6	26,891.00	3.158204
10	9	26,038.00	3.058024
11	12	25,535.00	2.998949

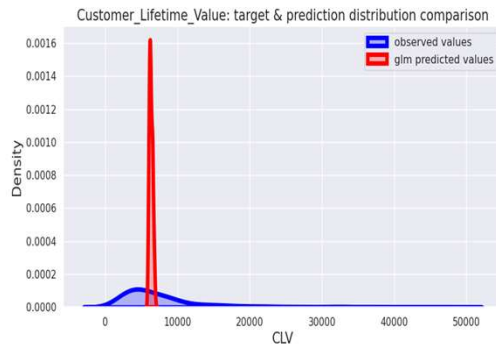


Which model can I use to predict CLV?

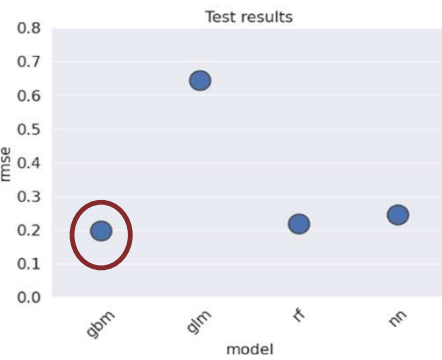
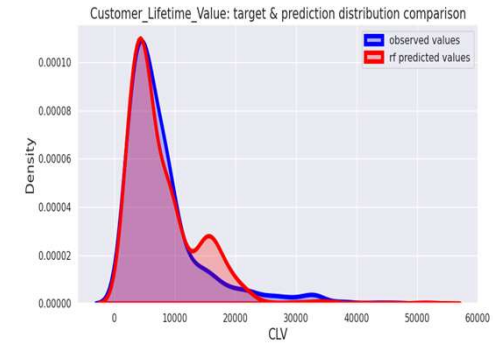
- Generalized Linear Model (GLM);
- Random Forest (RF);
- Gradient Boosting Machine (GBM);
- Neural Networks (NN)



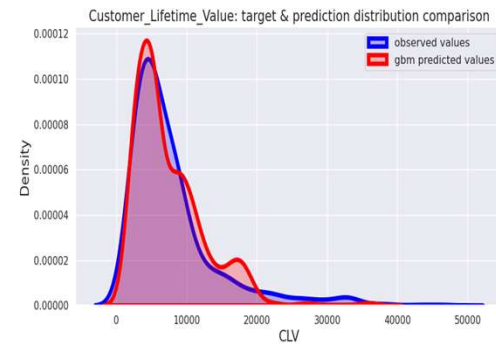
GLM



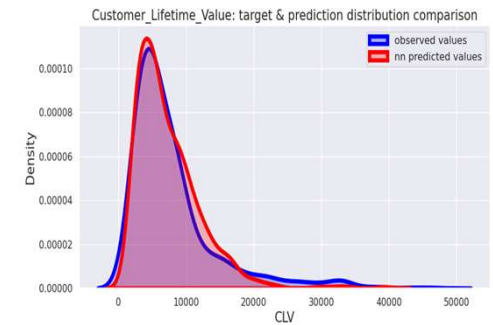
RF



GBM

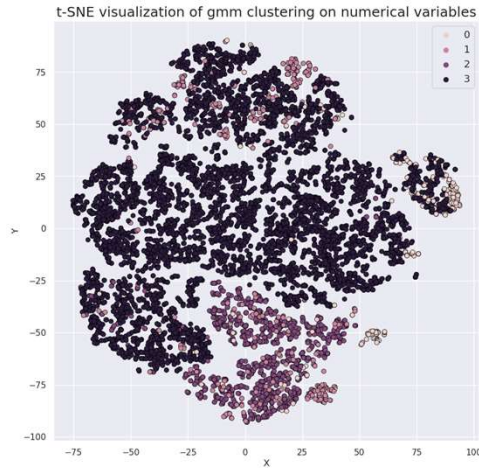


NN

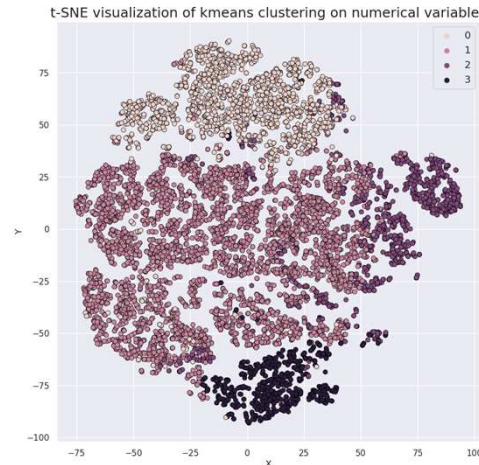


Clustering as Portfolio Analysis

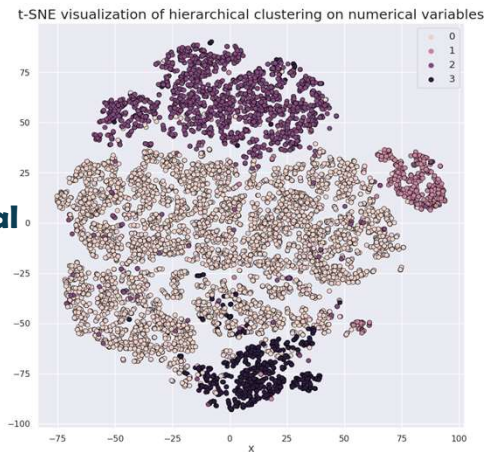
t-SNE with GMM



t-SNE with K-Means



t-SNE with Hierarchical Clustering



Improve portfolio analysis by the combination of t-SNE and Clustering tools and data visualization become a powerful tool leading portfolio queries

Query examples from k-means clustering



	kmeans_cluster	Monthly_Premium_Auto	PERCENTAGE
0	1	446,140.00	52.396751
1	2	182,378.00	21.419318
2	0	155,063.00	18.211318
3	3	67,884.00	7.972612

	kmeans_cluster	Total_Claim_Amount	PERCENTAGE
0	1	1,871,219.44	47.193821
1	2	1,129,102.60	28.476973
2	0	662,559.65	16.710344
3	3	302,085.36	7.618862

Cross-Selling Assessment

Go to the article



agentenews.com

What is Cross-Selling?
Cross-selling involves selling additional items related or complementary to a previously purchased item

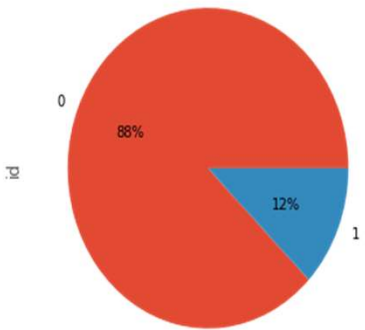
Use case: Insurance Cross-Sell Data Set

	count	mean	std	min	25%	50%	75%	max
id	381109.0	190555.000000	110016.836208	1.0	95278.0	190555.0	285832.0	381109.0
Age	381109.0	38.822584	15.511611	20.0	25.0	36.0	49.0	85.0
Annual_Premium	381109.0	30564.389581	17213.155057	2630.0	24405.0	31669.0	39400.0	540165.0
Vintage	381109.0	154.347397	83.671304	10.0	82.0	154.0	227.0	299.0

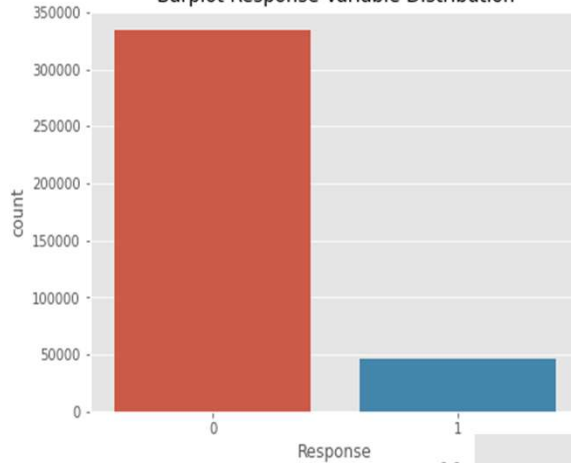
	count	unique	top	freq
Gender	381109	2	Male	206089
Driving_License	381109	2	1	380297
Region_Code	381109.0	53.0	28.0	106415.0
Previously_Insured	381109	2	0	206481
Vehicle_Age	381109	3	1-2 Year	200316
Vehicle_Damage	381109	2	Yes	192413
Policy_Sales_Channel	381109.0	155.0	152.0	134784.0
Response	381109	2	0	334399

Pay attention in the prediction...

Pie Response Variable Distribution



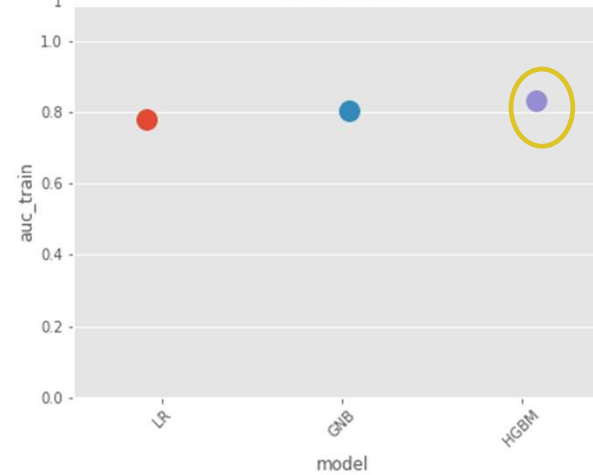
Barplot Response Variable Distribution



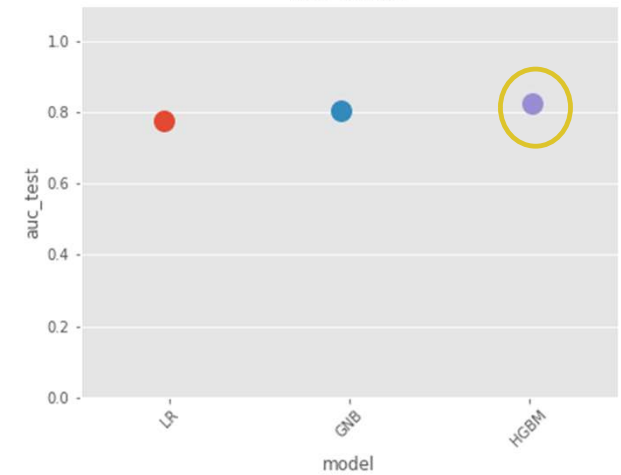
In the choice of the best performance model, remember to check if the model is calibrated

	model	auc_train	auc_test
0	LR	0.779368	0.778922
0	GNB	0.805083	0.803694
0	HGBM	0.832758	0.826070

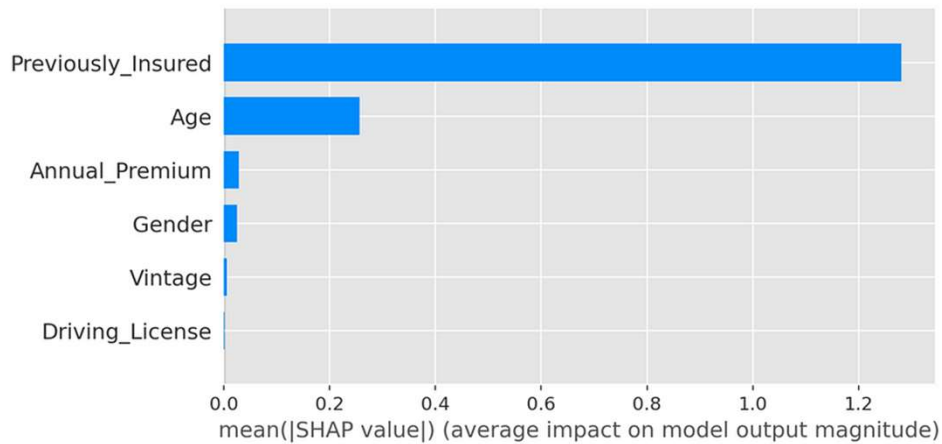
Train results



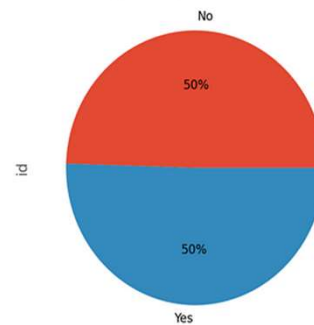
Test results



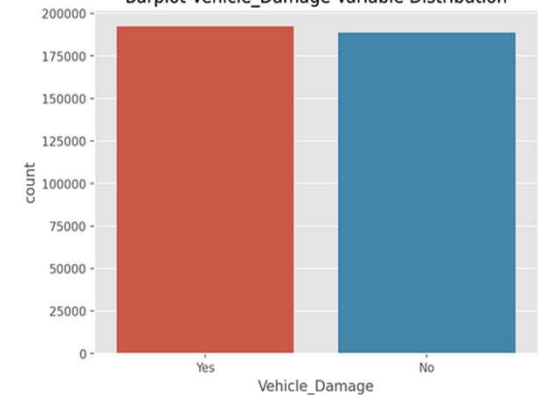
Look on other features than feature importance..



Pie Vehicle_Damage Variable Distribution



Barplot Vehicle_Damage Variable Distribution



Feature importance, say which are the most important features in the classification process, but take care about other features as «vehicle_damage» that can have an impact from the actuarial point of view as a possible risk factor in building the vehicle insurance tariff.

Customer Sentiment Analysis

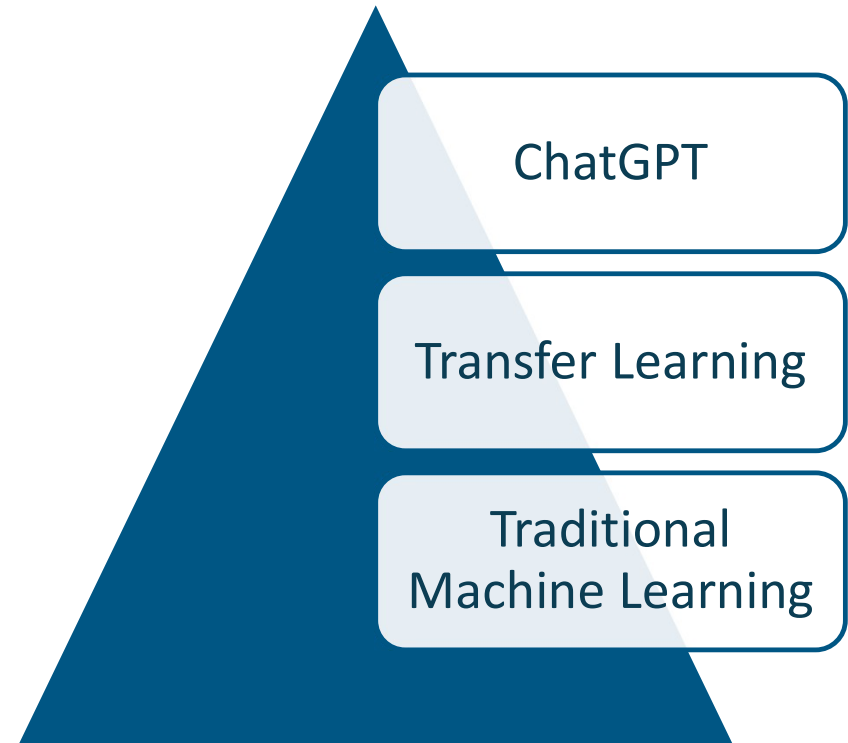
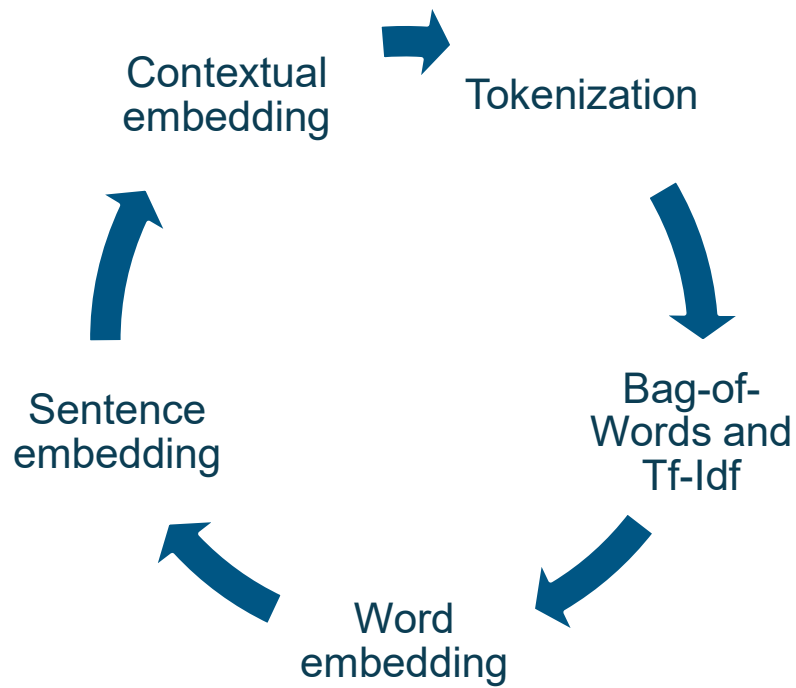


Sentiment Analysis

What is Sentiment Analysis?
Sentiment analysis is a technique that allows to extract users' opinions, emotions, attitudes expressed by customers from unstructured text data.

blog.reputationx.com

NLP World



blog.reputationx.com



Benefit of Customer Sentiment Analysis

- Customer sentiment analysis can help insurance companies to better understand customer perceptions of their products and services, and make informed decisions about product design, pricing, marketing, and customer service**
- Customer sentiment analysis can also help insurance companies to identify customer needs, preferences, expectations, and satisfaction levels, and to improve customer loyalty and retention**

Conclusion

- **Actuarial skills are valuable and versatile in marketing**
- **They can help marketers to segment customers, price products and services, and predict outcomes**
- **They can also help marketers to measure and manage risks, uncertainties and opportunities**
- **Actuaries need to improve their communication ability, creativity and strategy thinking towards innovative solutions, flexibility and never stop learning**

Thank you!

Contact details

Email: claudio.giancaterino@intesasanpaolovita.it

Linkedin: <https://www.linkedin.com/in/claودیods/>