**Details**: Your task is to investigate certain topics related to credit scoring. For this, you must write it as a form of an essay using APA style/format.

***It is very important TO CITE AUTHORS, if possible, for each written paragraph.***

**Structure of the chapter**

**Chapter name:** Credit Scoring review

**Sub-chapters:**

* best practises in credit scoring
* classical scoring approach
* drawbacks of the classical approach
* comparison between classical and modern credit scoring approach
* Obstacles present in classical and modern credit scoring.
* Negative betas problem in credit scoring.

**Specific information to emphasise in some subchapters:**

1. Information to be added to the **drawbacks of those classical** approaches subchapter:

* Explain what is [“weight of evidence” (WOE) transformation and provide an example in credit scoring](https://www.listendata.com/2015/03/weight-of-evidence-woe-and-information.html). Benefits and [drawbacks](https://stats.stackexchange.com/questions/189568/replacing-variabldrawbackses-by-woe-weight-of-evidence-in-logistic-regression) of using these techniques
* Explain that on classical approach we are not able to model nonlinear relationships between variables.
* problem with selecting features: Explain that variables cannot be selected automatically using statistical tests or Machine learning techniques, they must be done manually.
* Explain that there are not interaction between variables in classical approach: if we observe negative betas we must then create an additional feature (beta1\*beta2) manually that represent this “nonlinear” relationship between variables.
* Explain that many of these things that are done manually in classical approach, can be also done automatically with machine learning. And slightly give an explanation to what is machine learning and its applications in the fintech sector, especially in credit scoring.

**Specific information to emphasise in some subchapters:**

1. Information to be added to the **the Negative betas problem sub-chapter:** This is related to negative betas in a regression-based problem: Y =

* You must explain that If we create logistic regression based on WOE (weight of evidence), and IF WE OBSERVE NEGATIVE BETAS, it means that we have sympson’s paradox in our data. It means that we have higher other interaction between variables. However, IF WE HAVE POSITIVE BETAS, it means that we don’t have any interaction between features.
* Also explain what is the sympson’s paradox.
* NOTE: If we have negative betas but WOE was positive slope it means that there are some interactions between features. (Check prof Kaszynski book)