

Weight Of Evidence (WOE)

همانطور که در مقاله «[فرآیند توسعه کارت امتیازی](#)» به طور مفصل شرح داده شد کارت امتیازی مجموعه‌ای از متغیرها، مشخصه‌ها و امتیازهای آن‌ها است که با توجه به هدف سازمان متقاضی اعتبارسنجی طراحی می‌شود. در فرآیند طراحی کارت امتیازی عوامل متنوعی از جمله جامعه مشتریان، شرایط کسب‌وکار و سیاست‌های توسعه بازار سازمان متقاضی در نظر گرفته می‌شود. به همین سبب بسیار محتمل است که متغیرهای معنی‌دار در کارت امتیازی یک سازمان با سازمان‌های دیگر متفاوت باشد.

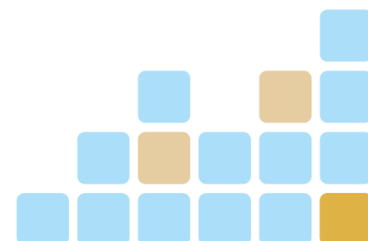
با توجه به هدف سازمان از اعتبارسنجی، متغیرهای در دسترس شناسایی شده و قدرت پیش‌بینی‌کنندگی هر کدام به طور جداگانه محاسبه می‌شود. سپس متغیرهای معنی‌دار، در کارت امتیازی به کار گرفته می‌شوند.

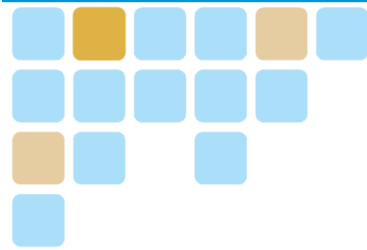
به منظور بررسی قدرت پیش‌بینی هر مشخصه (Attribute) در متغیر مورد نظر از مفهومی به نام Weight Of Evidence (WOE) استفاده می‌شود. دامنه و روند WOE مشخصه در متغیر مورد نظر گروه‌بندی می‌شوند.

در حقیقت WOE نشان‌دهنده قدرت هر مشخصه (مشخصه‌های گروه‌بندی‌شده) در جداسازی مشتریان خوش حساب و بد حساب است. WOE معیار تفاوت بین نسبت خوش حساب‌ها و بد حساب‌ها در هر مشخصه است. این شاخص بر اساس لگاریتم احتمالات محاسبه می‌شود، که معیار احتمال خوش رفتاری مشتریان را نشان می‌دهد.

$$(1) \quad \text{(توزیع بد رفتارها)} / \text{(توزیع خوش رفتارها)}$$

یکی از روش‌های محاسبه WOE استفاده از فرمول زیر است:





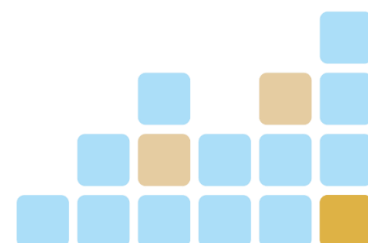
$$\left[\ln \left[\frac{(\text{توزیع خوش رفتارها})}{\text{توزیع بد رفتارها}} \right] \right] * 100 \quad (2)$$

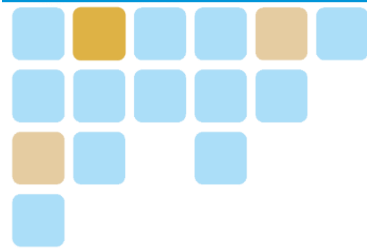
اعداد منفی به دست آمده از فرمول فوق، نشان دهنده بد رفتاری بیشتر نسبت به خوش رفتاری مشتریان است. به منظور فهم بهتر این معیار به مثال زیر توجه کنید. عملکرد مشتریان، در بازه‌های مختلف سنی در جدول ۱ نمایش داده شده است.

جدول ۱. بررسی رفتار متقاضیان در مشخصه‌های متغیر سن

سن	تعداد	توزیع کل	تعداد خوش رفتارها	توزیع خوش رفتارها	تعداد بد رفتارها	توزیع بد رفتارها	(WOE)
غیر دسترس	۱۰۰۰	۲,۵٪	۸۶۰	۲,۳۸	۱۴۰	۳,۶۵٪	-۴۲,۷۱۹
۱۸-۲۲	۴۰۰۰	۱۰٪	۳۰۴۰	۸,۴۱٪	۹۶۰	۲۵,۰۰٪	-۱۰۸,۹۸۰
۲۳-۲۶	۶۰۰۰	۱۵٪	۴۹۲۰	۱۳,۶۱٪	۱۰۸۰	۲۸,۱۳٪	-۷۲,۶۱۳
۲۷-۲۹	۹۰۰۰	۲۲,۵٪	۸۱۰۰	۲۲,۴۰٪	۹۰۰	۲۳,۴۴٪	-۴,۵۲۶
۳۰-۳۵	۱۰۰۰۰	۲۵٪	۹۵۰۰	۲۶,۲۷٪	۵۰۰	۱۳,۰۲٪	۷۰,۱۹۶
۳۵-۴۴	۷۰۰۰	۱۷,۵٪	۶۸۰۰	۱۸,۸۱٪	۲۰۰	۵,۲۱٪	۱۲۸,۳۸۸
بالای ۴۴	۳۰۰۰	۷,۵٪	۲۹۴۰	۸,۱۳٪	۶۰	۱,۵۶٪	۱۶۴,۹۳۴
مجموع	۴۰۰۰۰	۱۰۰٪	۳۶۱۶۰	۱۰۰٪	۳۸۴۰	۱۰۰٪	

در جدول فوق ۱۵٪ از کل جمعیت در بازه ۲۳ تا ۲۶ سال قرار دارند که ۱۳,۶۱٪ مشتریان خوش رفتار و ۲۸,۱۳٪ مشتریان بد رفتار از این بازه هستند. بر طبق رابطه ۱ داریم:



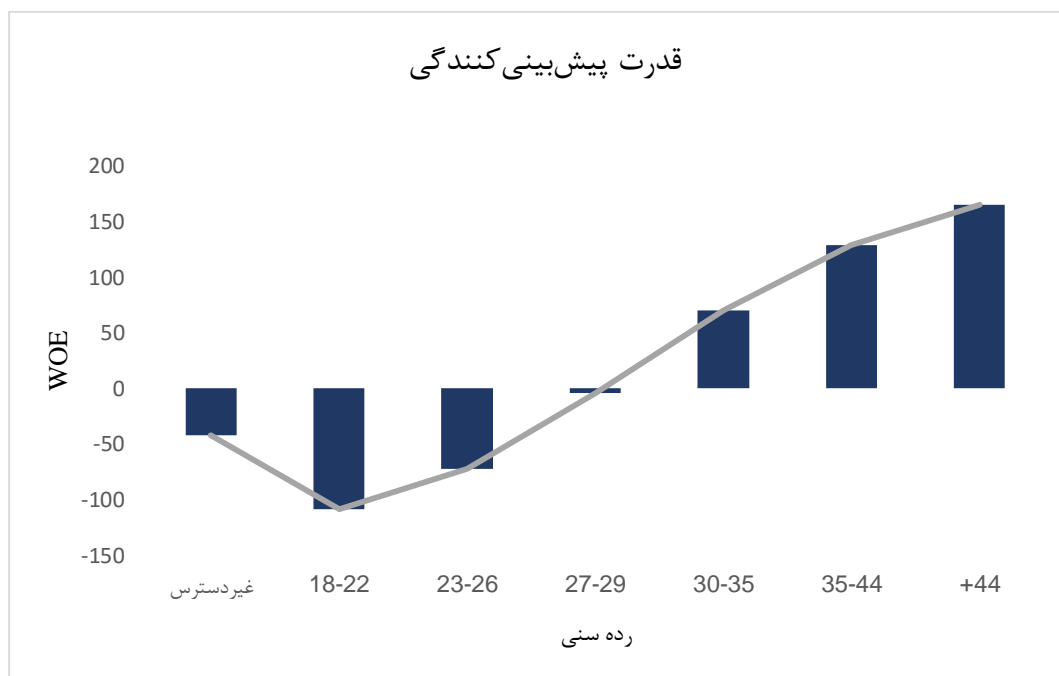


$$\frac{13.61}{28.13} = 0.48$$

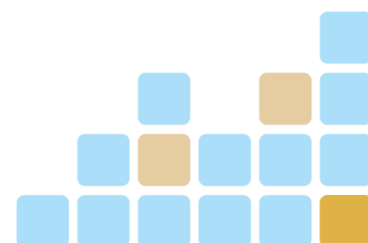
بر این اساس به احتمال ۰,۴۸ مشتریان در این بازه، خوش رفتار خواهند بود. WOE نیز از رابطه ۲ به دست آمده است:

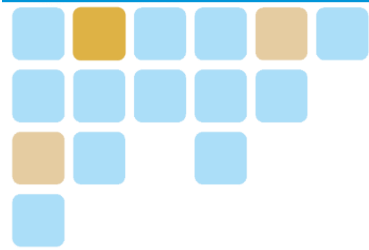
$$\ln\left(\frac{0.1361}{0.2813}\right) * 100 = -72.613$$

با توجه به اینکه WOE منفی شده است، می توان دریافت که رفتار بیشتر مشتریان در این بازه بد خواهد بود. در کارت امتیازی می بایست، قدرت مشخصه ها از یک نظم منطقی برخوردار باشد. منظور از نظم منطقی این است که WOE در نمودار سنجش قدرت پیش بینی کنندگی از روند خاص و معنی داری پیروی نماید. به منظور درک بهتر، توزیع وزن متغیر سن افراد مثال قبل در شکل ۱ نمایش داده شده است.



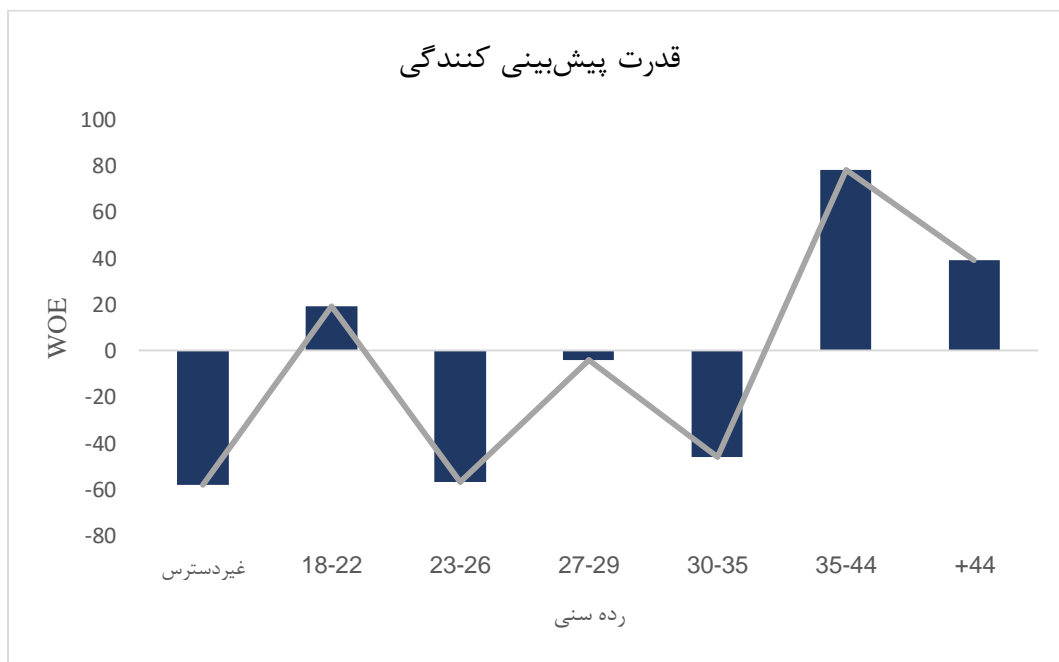
شکل ۱: روند منطقی WOE برای متغیر سن



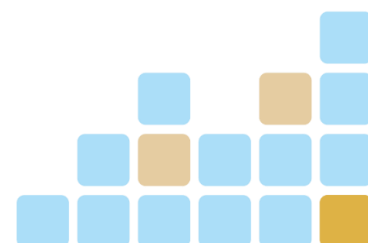


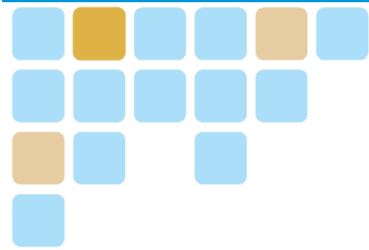
در شکل فوق اگر داده‌هایی را که در دسترس نیستند؛ نادیده بگیریم، بقیه گروه‌ها رابطه منطقی با WOE دارند. به عبارت دیگر، رابطه منطقی بین مشخصه‌های متغیر سن و نسبت مشتریان بدرفتار وجود دارد. به این صورت که در این مثال، پذیرش افراد جوان‌تر ریسک بیشتری به همراه دارد. ایجاد چنین رابطه منطقی‌ای به منظور گروه‌بندی متغیرها (تعیین مشخصه‌ها) یکی از اهداف تجزیه و تحلیل متغیرهای اولیه است.

شکل ۲ مثالی از یک روند غیر منطقی میان WOE و متغیر سن را نشان می‌دهد. همانگونه که در شکل مشخص است، در این مثال متغیر سن ضعیف است و هیچ رابطه منطقی بین این متغیر و رفتار متقاضیان (خوش حساب و بدحساب) وجود ندارد.

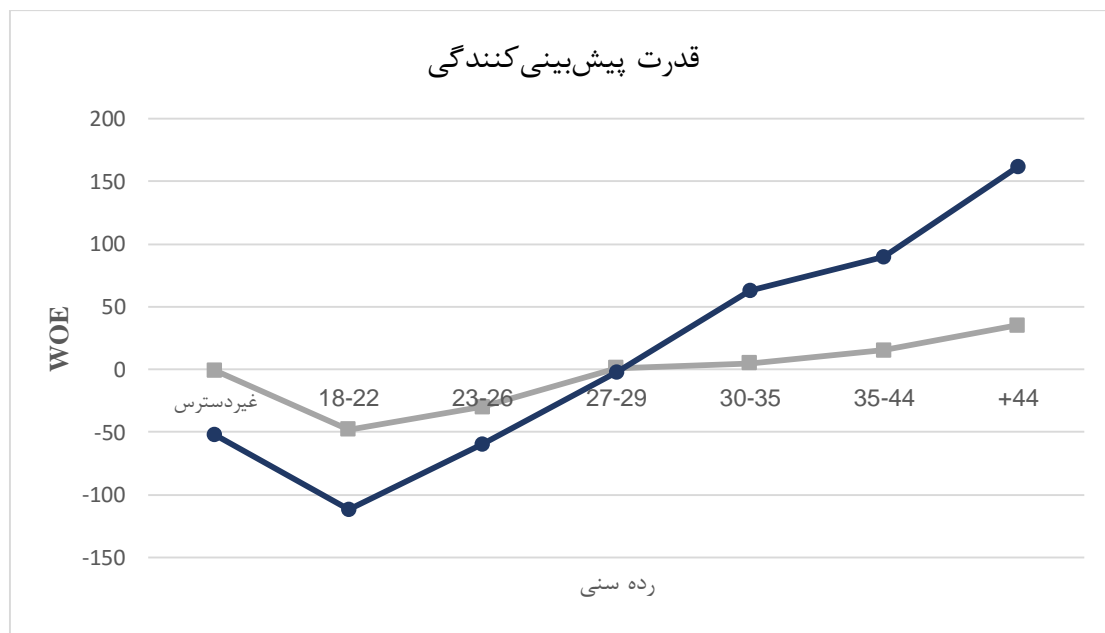


شکل ۲. روند غیر منطقی WOE برای متغیر سن





در شکل ۳ دو سری داده سن بر روی یک نمودار ترسیم شده است که هر دو دارای روند منطقی با WOE هستند. اما نمودار با شیب تندتر (نقاط دایره‌ای شکل) پیش‌بینی قوی‌تری را در خصوص رابطه میان سن و چگونگی عملکرد نشان می‌دهند.



شکل ۳. مقایسه WOE برای دو سری داده

منبع

Siddiqi, N. (2017). *Credit Risk Scorecards: Developing and Implementing Intelligent Credit Scoring (2nd ed)*. John Wiley & Sons, Inc.

ما را در شبکه‌های اجتماعی با [@icbsco](https://www.icbsco.ir) دنبال کنید

