چالش های ماتریس هم رخدادی:

1. اگر کانتکست (جایی که ماتریس هم رخدادی رو از روی اون می سازیم) را کل مخزن متن در نظر بگیریم، ابعاد مجموعه داده بسیار بالا می شود. مثلا در صورتی که V توکن (Term) منحصر بفرد داشته باشیم، طول ماتریس هم رخدادی ما برابر |V|\*|V|می شود. مثلا اگر V=1000 باشد ماتریس ما 1000 در 1000 می شود. هر توکن باید با 999 توکن بعدی خودش بررسی شود.
2. تجزیه مخزن متن به k زیر سند یا کانتکست. در این روش ما به اندازه دلخواه اندازه سند را تغییر می دهیم. مشکلی که در این روش وجود دارد برابر نبودن توکن های موجود در هر سند می باشد. که موجب می شود ابعاد کل توکن ها در مجموعه داده برابر نباشد. راه حل این است که تعداد توکن های متن را به گونه ای تهیه کنیم تا هر کانتکست تعداد توکن های برابری داشته باشد. مثلا متن با 1000 توکن به 10 توکن 100 تایی تقسیم می شود. این روش اعتبار لازم را ندارد.
3. اگر متن به 10 قسمت 100 تایی تقسیم شود، ممکن است در مرز این فولد ها توکن های معنا دار شبیه به هم قرار داشته باشن که با تجزیه کردن متن به این فولد ها این توکن ها از هم جدا می شوند و در کانتکست خودشان بررسی می شوند.

ماتری هم رخدادی اولیه خروجی مناسب ندارد.

ماتریس هم رخدادی به صورت باینری هم خروجی مناسب ندارد.

کد گزاری روش one hot encoding خروجی مناسب دارد.