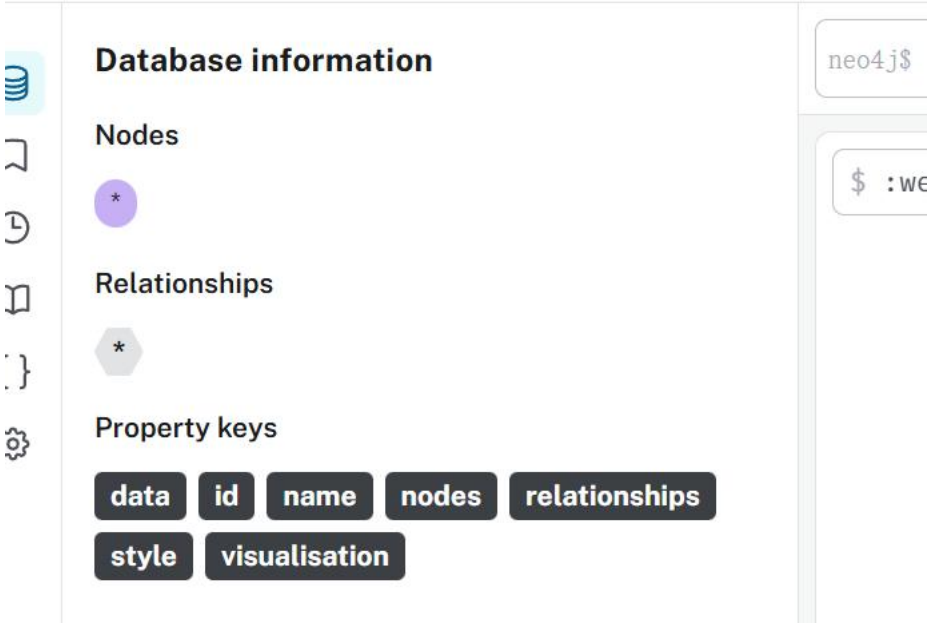
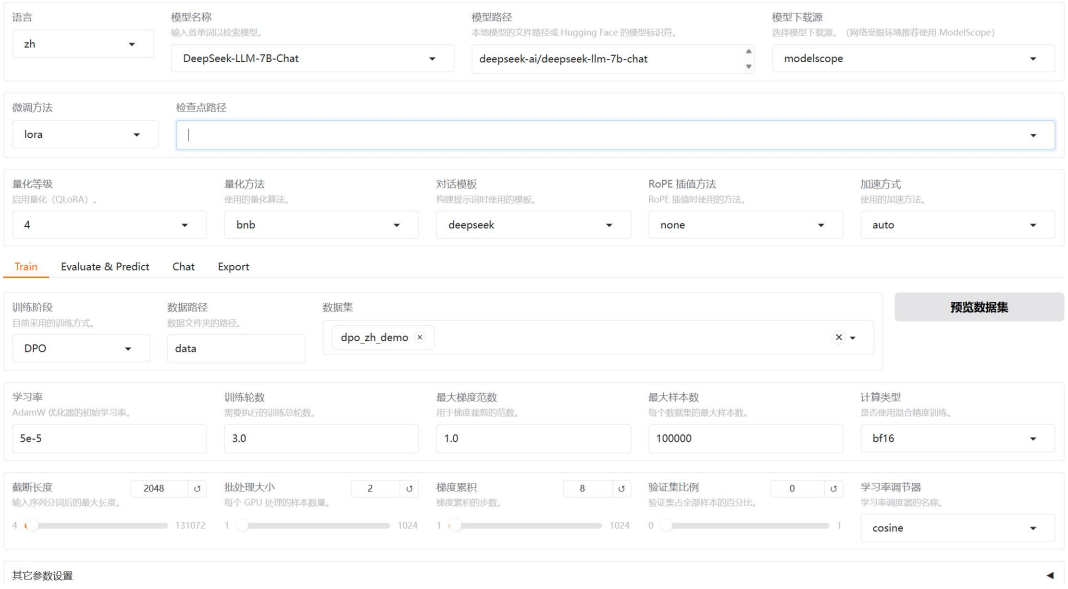
"Transformer 是一种深度学习模型，它采用自注意力机制（Self-Attention）来处理序列数据。它主要由编码器（Encoder）和解码器（Decoder）组成。BERT 和
GPT 都是基于 Transformer 的架构。



"



模型	类型
GPT-4	生成式
BERT	判别式