**微信 林白白白白 2025-09-11 22-34-33**

哦，不，我我现在做的一个已经做成的，是把一个人或者是比如说几个博主，他们的视频的文本直接爬起来。  
然后呢，我要分析的两个要素，第一个是流量，第二是需求，你不用把它全部爬去，你你预期全部爬出来，还不如直接直接把一个呃视频直接抄下来。  
或者是把几个视频轴承一个视频抽下来。  
嗯嗯嗯，你不用怕去做这么大的工程样，他也他也不知道，因为里面有些视频，他有些视频他他都不说话，有些视频没有工程量，没有任何工程量，1秒钟解决，我不得攻，我不得工工部工程量的问题就是。  
哎，你他妈全部爬下来，没意义，怎么不呃，20可能是因为我我的理论假设第一条流量就是。  
呃，怎么形成这个一个对于流量的网站，那我的理解就是它只有看得足够多了，它才有网感。那我就要模拟说让ai看着足够多的这样的一个过程，让它形成网感。  
形成网杆之后，然后呢我就要去再让它总结用户画像，然后它对于用户画像的概括，就是总结根据文本倒推用户的需求。  
然后再让它这个过程当中，我其实是这样这样子，你可以这样理解就是。  
不是你，你让他去搞王干之前，你先让他去明白平台的退路机制吧，我操。  
呃，不不是不是是这样子，是是是我我的意思是咱们先提取一个最关键的点，什么叫就是最关键的点，就是我就抓需求和流量。我认为解决了用户的需求。然后还有一些参数，我不知道是什么。  
我只是需要把这些参数给找出来。如果这些参数我们调对了，那这个流量就会出来。所以我现在不知道是什么，只能让ai自己去假设。然后我我在做的事情就是不要不用把，就是我输入的东西是什么？不用管，不用管，输出的东西是流量高的东西就可以了。然后ai跟流量机制划上等号，让ai去模拟流量机制，对于我们输入的文章进行一个评级。  
然后再去输出流量。就是那那现在我要怎么做这个事情，就是我把大量的文章给到他，让他先模拟先学习网感。  
然后学习网感之后，我让他分析用户的需求。因为我的前提假设是根据需求的这些文章里筛筛选过没啊，不是你还你记得听我说完，不是说因为这一步其实其实是这样子的，是我先得去理解一下，他是就是我我的我的前假设是这样的，是只有完全符合了用户需求的。  
文章或者内容它才能够去获得，就是需求可能是一个点，流量可能是一个点。那么它那个东西不需要删。因为我要的是它的判断的需求。因为它根据这个大量的文章就可能能看到。就是比如说你是一个做立体的，然后你的账号当中他就能看得出来你的用受众是什么样子，所以我是。  
把黑箱给到ai，把能看得到的给到我们自己，让ai通过看得到的去模拟这个东西。  
然后我不用关注太多的细节，我只是我只需要去调整他对于这个事情的理解。  
然后然后我对照了一下他的理解，你看他的理解还是很准的。  
所以另外一个点我也没看，等一下，另外一个点他就要测试了。  
因为他跟我讲到的说人对于发展的需求其实是排在中间的，其实并不是很强烈的。他跟我说这一点，所以其实ai在因为我我自己的身份证。  
思考逻辑是这个样子，那个网杆跟大量的阅读数据有关，就是跟你网上冲浪有关。那么网上冲浪它显然是比你了解流量机制更加的重要。  
所以我给他的不是我说的，我说的是平台机制，不是流量机制。  
你你指的是具体是指的哪1块平台，它的推流机制的变化。就比如说现在抖音和b站本身已经在变化了，我能够感受得到。但是但是具体的我不太能够描述的清楚，我要需要让他去搞一下。所以所以这个问题其实属于一个很细分的问题。我现在更多的是考虑，这不是这这不是很细分问题，这这个是要要命的问题哦，我知道这个可以解决啊。我的意思是先把那其实解决你这个问题也可以啊。  
就是我的意思是就看你选择哪些，我们先把模型定出来，看怎么调参数了。  
举个例子，因为我并不知道我这边要进隧道了，就是我我并不知道的是流量到底是怎么产生的。但是我知道的是什么？知道的是用户需求和用户画像，这个是可以验证的。那么我就通过用户需求，还有用户画像。  
就是用所谓的用户痛点，然后跟这个内容之间的关系。  
然后让他去猜测流量的高低。  
看他猜的准不准，如果他猜的不准的话，我再问他再去调整原因，就是跟他去分析嘛，然后这个时候我就慢慢地去做测试。  
因为过去过去，为什么没有人做这个东西？因为它。  
第一个是大量爬取文本，我们只能爬视频，爬不了文本，就是很多人他是爬不了文本的。  
而且他不能做到批量爬文本。那我现在有一个方法可以批量爬文本，但是为什么就算有了批量爬文本，别人还做不到呢？  
因为大家以为的数据库，它可能只有几十份文档。  
而我这样批量搞，他会搞到上千份文档。我认为只有上千份文档，ai才能看到里面的规律。  
然后它有可能能洞察到我们洞察不到的东西。  
那我的前提假设是，既然这个流量机制是。  
机器做起的就是代码的。那么反过来来说，让ai去模拟这个流量机制，我们只用掌控输入和输出去微调，就这种理论上的微调，这跟他们那个微调还不一样。  
然后让他每1次把他的思考过程都不是每次的思考过程都写成一份文档。我们把这个文档当成一个point，也就是。  
我只知道正确的答案是什么，还有我要的是什么，然后让他不断地去调整中间过程，以至于达到我们想要的。  
所以可能我下一步要做的就是因为我现在只选一个细分赛道，我选的非常细。  
所以只有选择细分赛区那个什么赛道啊，龙卷我记得。  
这个这个我爬的是龙镇啊，我知道你说的你选的什么赛道，18~25岁大学生大学生认知。  
这个赛道细吗？18~25岁这个归了啊，我的意思是我限定限定了年龄，18~25。  
啊，这个不细啊，我靠，那好，那你帮我再调一下，我负责帮你，我帮负责帮你调整。没？我先跑一个想法嘛。然后我还在想，诶，不是不是我的意思是说你。  
细分赛道，他他他是一个诅咒型的，你你越冲着细分赛道去，你越做越做不起来，我就感觉是可以ok这个没有问题。所以我就我就冲的是我就冲着这个18~25岁大学生认知这1块走的。  
然后我就把这类博主的内容让ai都看一遍，让他形成网感。  
  
然后下一步就是让他去分析这个画像分析可能会报的选题。  
然后我就我就看他为什么分析这个选题会报他的逻辑。如果这个真爆了，那我就证明他的逻辑是正确的，如果这个没报他就错了，不不？  
你不能你不能用这个的话，你不能用二分法去判断的。我靠，我不是二分法，其实就是这样子的，是我我让他判断两个面，他要判断两个维度。第一个维度是。  
他判断这个视频是不是会爆掉，这是第一个位置。第二个是他要判断这个视频没有流量，他是要判断两个为主，第一个是有流量的视频，第二个是没有流量的视频。  
但是他其实做到了第三个维度，就是他又判断了说有些人蹭热点的事情。  
说明他对这个分类是比较轻的，但是我的意思是这些都不重要，重要的是它量大。  
它量大是重要的。所以它量大的时候，我们不需要它生成100张100。  
两篇文章当中都是正确的，我们只需要它大概率正确就可以了。我做的这个东西它只是需要粗略的，所以还最终还是要你来审稿。  
我只是现在提了一下，因为过去做不了这个事的原因是在于呃，我用的是扣的，扣的靠的扣的，它可以直接的在命令行当中运行，它可以。  
它可以有个巨大的数据库，就是它可以瞬间读取上百篇文章，上千篇文章都不是问题。但是国内没有人能够想得到这个他用国内用克劳的扣子用的较多都是编程。而我把这个事情也理解为编程就是  
我把grant变成了代码，这个代码是用来执行最后的流量的。  
所以说流量输出的结果就是流量中间的，我们可以调整。  
所以我现在的问题就在于是要去建立一套这样的关于自媒体的这样就是四个关键词，然后它的输出模式，然后这样去调这个东西，只要他。  
10篇当中有三篇是爆款的，我觉得10篇当中有一篇是爆款的，就已经够了。  
然后我怎么样判断他预测的准不准。所以要建立一个预测的标准，就我可能让他判断。就比如说我让他判断两个维度，我给他了500篇文章，他只用判断判断500篇文章当中流量最高的10篇文章和流量最低的10篇文章。  
结果发现了个问题，就是他妈对流量低，判断得特别准。  
他跟我说，这10个流量特别低的，全都是1万以下的，就很准他对流量高判断的。  
就是就是你你能反向通过，就如果他他慢慢特别准的时候，说明他的思考逻辑是正确的。  
所以我我不是真的给他流量，让他学习，而是通过后果测试来反映他学了没学。  
这个东西之前的人的思维是完全不一样的。我他妈之前用ai的时候，我他妈给流量数据给他，让他刷。  
这肯定算不出来呀，然后我验证又没法验证了，但是这样这1次的话，我就通过。  
先让他思考用户的画像，然后再让他思考用户的需求。  
然后再去让他推测哪一篇文章可能报哪篇文章爆不了，说明原因。  
然后他跟我说，我做，他他这样去做，我就觉得这事儿可以搞。  
你确定大概都说了啊，准确率你大概多少？你那边算出来的话，你是不是算出来的？他对我我只是大概我今天晚上初步测试了一下，他对于非爆款的文章识别率，就是他给我了10个全都是1万以下的。  
所以我等会儿等会儿你呃，你对爆款的判断，一判判断播放量是多高。  
我的意思是不用判断多高这个东西他你知道为什么吗？因为不是不是不是我说的，我不是你，你听我说我说的是你说的爆款，因为我不知道你刚才说的爆款的定义是什么，所以我现在要要给我要问你一下，呃，是这样子的，不然我听不懂你在说什么。  
这是相对的，就是ai，他肯定能把握的一个东西是。  
比如说他把握到了一个关键词，这个关键词关键数据越高的话，这个流这个文章流量越高。但是这个文章流量具体有多高，他是没办法去判断的。它只能说比如说因为为什么呢？因为它只能判断这个东西流量高，而不是具体流量是多少，我们也不需要它判断具体流量是多少，我们只需要让它判断流量高低的准确度越来越高就可以了。  
你懂吗？嗯，所以你刚刚说那个就很好，很有意思，你给他一个一个10万播放的一个1百万播放，看他能判断出来哪个高吗？  
然后如果他判断错了，然后他的喷气逻辑是什么？然后直到我们微调到所有的你你说的爆款狮子这个相相对流量，对不对？他只能判断相对流量，他不能判断具体流量，因为具体流量它几乎就很复杂了。  
嗯，就没必要，我的意思是，我只是想把它训练成一个。  
他都无所谓。  
流量是什么就可以了。然后如果我的前提甲是这样的，既然安先生能够教会别人懂这个东西，那么ai懂的不是不应该比他更快吗？  
嗯，只是因为那里保安现在是八八上也没有，我回头准备扒一下。我这边我后续把这个我的那你那里在拼多多上面买一下，盗版就有了，我以为他的盗版错了，我的。  
我对，然后我啊不行，安逸，现在不行，他商业化不行。对对我我的意思是我只需要找到一个能够供我ai训练的合合格的参数，这个参数找到了我这个工作流就可以自动化了。  
所以现在的问题你可以找到机吧，我靠啊七夕我靠，可以可以都可以。所以我的我的想法是，  
卧槽我都我都我都怀疑这个事儿是不可能的，但是他前提理论上是可以的。  
先揍了再说，我靠。  
对呀，我就觉得这个事很兴奋，所以我感觉很有意思。我现在已经发给你了。他第1次的分析。  
因为我要的什么东西，我要的就是一个他对于我给他的大量的文章的。  
一个分析，然后我不管他分析了什么内容，然后呢，我要的输出的结果是他的判定结果。  
如果证明如果能证明他的判定结果是对的，那就能证明他的分析过程是对的，就是我的意思是我只能证明这个。  
嗯，所以要的就是这一个东西。好，ok你可以看一下他写的东西，我觉得  
我看了我看了不够不够，等一下，你先我先点一份外卖，我就没事。我已经说完了，你你看我明天再去测试。哎，我也在想，我明天到底要不要去，我要去的话，就要跟太阳请个假。  
你们想去就去呗，你哦，他让他说他们大概会多久离开哟。  
没有没有没有啊，没有，我也没有想到杨毅主要是看你们要不要过哦，你们要不要去，然后然后你们要去的话，他才过去，然后那这样肯定去啊，他他怎么办？是下午下午左右的样子，下午下午2点到4点的样子的样子的理由了，我约了杨钰莹好的。  
哎，你这边能走你的那个去吗？就就是你是不是还有一个你还有两个名额，对吧？  
我还有一个哦不对，我他妈姓杨义过去哦，算了，你一个加一个嘛，两个嘛，还有还有我一个是他妈，那我就跟太阳，我找到理由了，我跟太阳的理由就是我要去跟杨一聊一些事儿。  
嗯，好的好嘞，不然不然牙没人给啊，你可以先看一下他这个分析，我觉得有点。  
因为我我做的这个事情，他方便测试了，他不用我写文章了，所以他。  
能自己测试。  
我重点看重的是这个你你倒也不知道写不写文章，你你需要做一个审核员。对呀，所以我觉得有意思的是这个呀。  
Ok没意思。唉，我觉得没啥意思。