

算命术是数据拟合吗？

每年回家过年都能听到算命相关的话题，今年又恰好看到复旦哲学系的王德峰教授讲的命理学的一个小故事，我更加感兴趣，算命术到底work不work？本着不迷信的原则，我们不能不做调查和实践就说它是假的，也不能不做调查就说它是真的，正好假期，就找来几本书来看一下。

事先声明，本文纯属业余随笔瞎写，本人五行缺杠，杠就是你对：)

首先需要明确的是，我们这里说的算命术或者叫命理学就是一个普普通通的技术，不涉及什么鬼神，也没什么神秘的，在过去中国的知识分子里懂命理学的非常普遍（比如红楼梦里作者就给元妃安排过八字），普通人看一些命理书（比如滴天髓、三命通会）就可以看明白。大多数人觉得神秘主要是因为里面的语言和术语距离现在比较久远，读起来比较晦涩。

本文分两部分，第一部分讲算命术的基本原理和发展历史，第二部分讲算命术作为一个拟合了大量数据的规则系统，可能是怎么开发出来的。

一、算命术简要原理及历史

作为在现代科学教育下熏陶长大的一代，对算命这套含有封建迷信色彩的东西第一感觉是排斥的，因此大部分人不太去主动了解它，所以这里先对它的基本原理和历史做一个简单的介绍。

四柱五行

现代民间流传最广的算命术就是四柱预测，俗称八字。四柱即年柱、月柱、日柱、时柱，表示出生的年月日時，每一柱有俩字，一个表示天干，一个表示地支。天干共有十个：甲乙丙丁戊己庚辛壬癸，地支有十二个：子丑寅卯辰巳午未申酉戌亥。天干地支都和五行对应，比如甲乙对应木、丙丁对应火，其中又有阴阳之分，比如甲是阳木，乙是阴木，丙是阳火，丁是阴火。通过分析八字对应的五行就可以推断一个人的命运。这里我们举一个简单的例子，假如一个人命属木（一个人的基本五行属性看日干，命属木说明他日柱的天干为甲或乙），按照五行相生相克的说法，水生木，金克木，那么如果四个地支里面有亥子（表示水）或者其余三个天干里面有壬癸（表示水）来生木，那就是好的，如果天干有丙丁（表示火，火克金，从而保护木），那也是好的；但如果天干是戊己（表示土，土克水，从而抑制木）或庚辛（表示金，金克木），地支是辰戌丑未申酉（辰戌丑未表示土，申酉表示金），那就是不好的了。按照一般命理书的说法，四柱流通，五行生化是好的，四柱缺陷，五行偏枯就是不好的。分析其他三柱和日柱（我）的相生相克关系，又衍生出五个概念：印星（生我）、官杀（克我）、比劫（同我）、食伤（我生）、财星（我克），每一对相生相克关系里面如果是同性的（阳阳、阴阴）就是偏（比如偏印、偏官），如果是异性的（阳阴、阴阳）就是正（比如正印、正官），总共十个称为十神。基于以上四柱五行十神的概念，找一个用神（有用之神，也叫喜神，通常是五行里缺的，或者能生我的，或者我克的；找到对的用神是算命术的根本所在），按照一些公式计算一下大运和流年，分析一下流年中能克我的年份（比如命属木，金克木，木年为庚、辛开头的

年份)就可能有危险,我能克的年份(比如命属木,木克土,土年为戊、己开头的年份)就可能发财。

可以看到,根据四柱来预测命运的原理好像也不难,甚至就用四柱还能预测祖上/幼年(看年柱)、父母/青年(看月柱)、家庭/成年(看日柱)、子女/晚年(看时柱)。再有就是神煞、纳音、齐格异局(比如有命书说四戊午牛逼),这些在《滴天髓》里被批评是“好事妄造,非关命理休咎”,因而到清朝时就已经不太流行神煞纳音了,主要看四柱五行。

算命术的历史

那么这套神奇的算命方法是什么时候形成,怎么形成的呢?命理书当然不是上帝送的,那也得是人写的。

首先,天干、地支最早就是用来记日的。四五千年前,用天干记日,用地支记月,十二个月用十二个地支表示没什么毛病,但是一个月三十天用十个天干表示那还有三天是同一个天干,所以就用天干和地支组合起来记日,后来又扩展到记年、记月、记时,天干和地支顺序地搭配下来一共是60个组合,所以60年循环一次,称为六十甲子。而阴阳五行就是中国古代的哲学概念,原本是唯物主义的,在中国古代很多学术领域(比如中医)都有对照,后来随着唯心主义出现开始被用到命理学里来了。可见,命理学里的概念和原理都不是独创的,因此是一门后生的学问。

中国的算命术大概是从北宋的徐子平奠定的基础,所以也叫子平术,到明清时候风靡全国,考试、经商、结婚、起名都得找人算一算,已经融入民俗。算命术在发明之后也不是静态的,而是动态发展的。在徐子平之前还曾有李虚中的三柱预测(没有时柱);还有按照受孕时间推算的,但受孕时间就不好说了,经常根据出生时间往前推十个月,但有早产有晚产,因此更不靠谱;过去还有以年柱而不是日柱为主推算命运的。所以说,算命术自发明之后就一直在调整。

总之,从历史发展上来看算命术是古人自发总结的一套预测方法,不是什么神谕,其中很多规律也受限于当时的认知水平,现在看来很荒谬,举个例子:命理书里都会提到五行和五方的关系,木向阳而生,东方是太阳升起的地方,所以对应木;金以肃杀为秉性,西方是太阳落山的地方,所以对应金。漂亮的孩子都学过地球自转,这对应关系显然扯淡。

二、一个可解释的规则推理系统

从CS/AI的角度来说,算命术就是一个部分可解释的规则推理系统,它的输入输出都是结构化的,运算由一些线性计算和规则定义,因此是完全可以程序化的。

算命术的输入是8维的,每个维度取值是10或12个离散的值,输出是一个人一生的命运(甚至不止,还包括祖上、父母、配偶、子女的)。注意这里的输入已经是预处理过后的了,原始输入可以认为是连续的时间,根据节气划分到对应离散值。算命术就是要完成输入到输出的一个映射,并且还是可解释的(运算都是线性的,规则也是可枚举的)。可根据前面我们说的,算命术的可解释性不能说是毫无道理,只能说是一点道理都没有。前面主要说了一些规则方面的,这里再补充一个运算方面的例子:在算起运时间的时候,假如是天干逢阳年生的男性,或者天干逢阴年生的女性,从本人生日的那天算起顺数,到下一个节止,以三天为一岁;反过来,假如是阳年生的女性或阴年生的男性,从本人生日的那

天起，倒数到上一个节止，也是三天为一岁。多下来的一天抵四个月，一个时辰抵十天。为什么这么算呢？说不清楚，那就只能是数据拟合的结果，那么拟合的准不准呢？

我只看了两天书，想必算不准，也不掌握很多八字数据，按照王德峰教授的说法，大概有六七成准确率，如果这里的准确率是严格的accuracy而不是像ROUGE、GLEU这种指标那这个效果还是non-trivial的。实际准度如何这里我们暂时无法得知，我们要考虑的是，假如真能达到这样的准确率，这么一套系统是怎么开发出来的呢？或者说，数据是怎么被收集和拟合的呢？

现在要收集训练一个机器学习系统的数据，我们一般优先利用已有的一些数据或人工标注一些数据，例如用已有的一些数据做预训练，用真实场景下的人工标注数据做一下微调。我不负责任地推测古人也是这么做的。

预训练（史书）

一个很好的早期数据来源就是史书，唐宋时期史记已成，其中帝王将相、贤人志士的出生和死亡年月日，生平事件有不少记载，对这些数据做一下预处理，用自己的五行假设建立一个框架，便可以找一些规律了。有两个佐证，一个是算命术发展早期，唐代的命理学是三柱预测，可能就是还收集不到很多八字，只能依靠史书来冷启动，而史书中记载的时间不会精确到时辰；再一个是许多命理书中确实有很多名人的命造例子，例如诸葛亮（辛酉丙申癸丑丁巳）、孔子（庚戌戊子庚子甲申）、关羽（戊午戊午戊午戊午）、张飞（癸亥癸亥癸亥癸亥）、吕洞宾（丙子癸巳辛巳癸巳）等等，但其真实性存疑，例如吕洞宾的出生年份命理书和史书就对不上，况且大多数人的出生时辰更是无从考证，很难不让人怀疑是不是拿着生平找八字。

微调（人工标注）

预训练之后，一部分算命先生就可以开始上岗了。虽然当时的计算方法跟后来的有许多不同（比如按年柱还是日柱，三柱还是四柱），训练数据也不太均衡（史书中主要是名人，没有普通人，实际应用面临domain adaptation的问题），但总归有一套在史书上拟合过的系统可以用了。在上岗之后，算命先生就担任了数据收集者的角色，不少算命先生把自己一生所见的命造记录下来。例如在包括《滴天髓》在内的很多命理书里都有大量的命例，里面有匿名的也有实名的（除了名人，基本还是做了些隐私保护的，比如李总兵、张太保、丁都督等等）。这些实践中收集的数据就更详细了，不仅有出生年月日，还有出生时辰，也因此才发展了四柱预测。有了新的数据，就可以fine-tune我们的算命系统了，例如，算出一个人神煞有天乙贵人，但是却发展不好，或者算出一个人神煞有亡神，但是却发展很好，那么就可以用空亡来打这个补丁，可以说一个好的或者坏的神煞遇到了空亡，所以其好的或者坏的作用没有体现。再加上各五行属性力量强弱、流年等等因素，算命这件事本身就变得非常灵活，大量的命例都能找到合适的解释。得益于自唐宋起很多年代算命先生的开源精神，终于在明清时期，算命广泛地推广开来了。实际上，直到民国还可以在命理著作里找到其收集的同时代的命造。发展到今天，算命系统复杂度至少足以拟合训练集，但至于泛化性如何，我们就无从得知了。

写在最后，很多中国人有一种根深蒂固的观念，就是迷信祖宗，越早的东西就越是好。金庸的武侠世界里也是这样，越早的武功越厉害，后来逐渐失传，降龙十八掌后来能打出十三掌就算牛逼了。而现实世界的发展通常不是这样的，通常是过去打出降龙十八掌是天下第一后来要能打三十六掌才能混口饭吃（就是卷）。因此，对过去前辈的智慧要尊重，但对于其成果也要祛魅，不能总是照抄照背，总归还是要发展的，比如拿三千张A100训练一个易1.0大模型。